DOCUMENT RESUME

ED 195 339 PS 011 849

AUTHOR	Ruopp, Richard: And Others
TITLE	Technical Appendices to the National Day Care Study: Effects Analyses. Final Report of the National Day Care Study. Volume IV-C.
INSTITUTION	Abt Associates, Inc., Cambridge, Mass.
SPONS AGENCY	Administration for Children, Youth, and Families (DHEW), Washington, D.C.
PUB DATE	Oct 80
CONTRACT	105-74-1100
NOTE	553p.: For related documents, see ED 131 928-930, ED 147 016, ED 152 114, ED 160 188, ED 168 706, ED 168 733, and PS 011 846-849.
EDRS PRICE	MF02/PC23 Plus Postage.
DESCRIPTORS	Child Caregivers: Classroom Environment: *Day Care: Day Care Centers: *Measures (Individuals): National Surveys: Observation: *Policy Formation: Preschool Children: *Research Methodology: *Test Results
IDENTIFIERS	*National Day Care Study

ABSTRACT

This volume, "Effects Analyses," the third of three Technical Appendices to the National Day Care Study (NDCS), presents in six individual reports the results of analyses investigating the relationships between policy variables, classroom processes and child outcomes. The volume begins with a presentation of the major effects analyses based on the two behavioral observation instruments used in the study ("The Adult Focus Observation" and "The Child-Focus Observation"), and then moves to a detailed treatment of the development and use of adjusted test score gains. Links among caregiver and child behavior, child test scores and other dependent measures are explored. Results of the Atlanta Public School (APS) controlled substudy and APS replication substudy are also reported. (Author/MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Final Report of the National Day Care Study
VOLUME IV-C

October 1980

TECHNICAL APPENDICES
TO THE NATIONAL
DAY CARE STUDY

EFFECTS ANALYSES

Papers by:

Barbara Dillon Goodson
David B. Connell
Robert Goodrich
Judith D. Singer
Nancy N. Goodrich

Prepared for:

Day Care Division
Administration for Children, Youth and Families
Office of Human Development Services
Department of Health, Education and Welfare

Allen N. Smith, Government Project Director

Prepared by:

Abt Associates Inc.
Cambridge, Massachusetts

Richard Ruopp,
National Day Care Study Director

Jeffrey Travers,
Associate Director

Nancy N. Goodrich,
Technical Coordinator

Contract No. 105-74-1100

PS011849

OVERVIEW OF NDCS FINAL REPORT VOLUMES

Results of the National Day Care Study and its major supporting study, The National Day Care Supply Study, is presented in a five-volume final report. Contents of these volumes are as follows:

Volume I

Children at the Center: Summary Findings and Policy Implications of the National Day Care Study presents in summary form the major findings and implications for federal day care policy of the National Day Care Study, a four-year study of the effects of regulatable center characteristics on the quality and cost of day care for preschoolers. Volume I serves both as a self-contained volume for the policy makers and as the foundation for the detailed presentation of results in Volumes II, III and IV. (Executive summaries of Supply Study findings and findings of an Infant/Toddler Study are included as appendices to Volume I.)

Volume II

Research Results of the National Day Care Study is a companion volume to *Children at the Center.* Volume II documents the analyses and results of the NDCS for the technical reader who seeks a more thorough understanding of the study from a research perspective. Volume II thus provides the quantitative support for the findings and policy conclusions reported in *Children at the Center.*

Volume III

Day Care Centers in the U.S.: A National Profile 1976–1977, the final report of the National Day Care Supply Study, is based on data gathered from a national random sample of over 3000 day care centers, stratified by state. Summary information is presented on characteristics of children and families served, center programs, staff, finances and regulatory compliance. Discussion of results is augmented by over 150 statistical tables.

Volume IV

Technical Appendices to the National Day Care Study is a compendium of technical papers supporting the most important conclusions of the study. These papers form the basis for the summaries in Volumes I and II. NDCS appendices are bound in three sections as follows.

Volume IV-A, *National Day Care Study Background Materials,* contains three papers, each of which establishes a distinctive context for the NDCS: a literature review focused on effects of group care and regulatable characteristics of the day care environment; case studies of the history and current practice of day care in the three NDCS sites (Atlanta, Detroit, Seattle); and a review of child development issues relevant to the NDCS from the perspective of black social scientists.

Volume IV-B, *National Day Care Study Measurement and Methods,* presents individual reports on a series of technical tasks supporting the principal analyses of the effects of key center characteristics on children. Among the topics covered are: analysis of alternative measures of classroom composition; psychometric analysis of the NDCS test battery; and analyses of several other more peripheral instruments used in the study. Also presented are results of a special survey of parents of subsidized children taken during Phase III, analyses of the impact on children of other center characteristics, such as physical space and program orientation, and econometric analyses.

Volume IV-C, *National Day Care Study Effects Analyses,* also a series of individual technical reports, begins with a presentation of the major effects analyses based on the two behavioral observation instruments, and then moves to a detailed treatment of the development and use of adjusted test score gains. The links among caregiver and child behavior, child test scores and other dependent measures are explored. Also detailed are results of the Atlanta Public School (APS) controlled substudy and APS replication substudy.

Volume V

National Day Care Study Documentation and Data gives a brief overview of NDCS data collection instruments and data files. Part A consists of the instruments themselves, including interview and data collection forms, observation systems and cognitive tests. Part B consists of data dictionaries; these describe every variable in the NDCS analytic data files. Part C provides codebooks for the data files. Parts B and C are available on computer tapes, which are readable independent of specific computer systems. Note that computer tapes are available only from Abt Associates.

Copies of the final report may be ordered from:

- EXECUTIVE SUMMARY (ONLY)
  Day Care Division
  Administration for Children, Youth and Families
  Office of Human Development Services
  Department of Health, Education and Welfare
  400 6th Street, S.W.
  Washington, D.C. 20024

- EXECUTIVE SUMMARY, Volumes I–IV
  ERIC Document Reproduction Service
  Computer Microfilm International
  P.O. Box 190
  Arlington, VA 22210

- EXECUTIVE SUMMARY, Volumes I–V
  Abt Associates Inc.
  55 Wheeler Street
  Cambridge, MA 02138

Earlier NDCS publications available from ERIC (hard copy or microfiche) are:

National Day Care Study First Annual Report, Volume I: An Overview of the Study [order number ED 131 928], *Volume II: Phase II Design* [order number ED 131 929], and *Volume III: Information Management and Data Collection Systems* [order number ED 131 930] (Cambridge, MA; Abt Associates, 1976).

National Day Care Study Second Annual Report [order number ED 147 016] (Cambridge, MA; Abt Associates, 1977).

National Day Care Study Preliminary Findings and their Implications [order number ED 152 114] (Cambridge, MA; Abt Associates, 1978).

PAPERS:

The Adult-Focus Observation
Effects Analysis
(1)

The Child-Focus Observation
Effects Analysis
(111)

Analysis of Test Score Growth
in the National Day Care Study
(209)

Classroom Process-Child
Outcome Analyses
(321)

The Atlanta Public Schools
Day Care Experiment
(397)

The Effects of Day Care in Eight
Atlanta Public Schools Day Care Centers
(433)

GLOSSARY

This glossary is intended as an aid to the reader. It is not an exhaustive dictionary of terminology relevant to the study or practice of day care, but rather a list of terms used throughout the volume which may be unfamiliar to the reader or which have special meanings for the purposes of the National Day Care Study.

An alphabetical list of terms enables the reader to find any item easily; numbers refer to the location of the term in the glossary itself, which is arranged by subject area to facilitate understanding of terms in relation to each other and in the context of this study. Subject areas are:

Classification of Day Care Services
Children and Staff
Classification of Day Care Centers
NDCS Independent Variables
NDCS Dependent Variables
Statistical Terminology

Alphabetical List of Terms

activity subgroup [42]
aide [17]
auspices [21, 25]
background variable [46]
caregiver [13]
caregiver/child ratio [44]
caregiver qualifications [45]
child outcome [51]
classroom composition [38]
classroom process [49]
core care [8]
correlation [59]
cost variables [54]
day care [1]
day care center [2]
dependent variable [47]
developmental outcomes [52]
effects [48]
family day care home [3]
FFP center [34]
full-time day care [6]
funding source [30,33]
generalizability of a measure [57]
generalizability of a sample [58]
group center [23]
group day care home [4]
independent center [22,26]
independent variable [36]
infant [12]
in-home day care [5]
lead caregiver [16]
lead teacher [15]
legal status [19]
multiple regression [61]

non-FFP center [35]
nonprofit center [24]
number of caregivers [39]
outcome [53]
parent-fee
part-time day care [7]
policy variable [37]
preschooler [10]
principal components analysis [62]
private center [28]
process [50]
profit center [20]
provider [18]
public center [29]
publicly funded center [32]
regression [60]
reliability [56]
sponsored center [27]
staff [14]
staff/child ratio [43]
staffing pattern [40]
supplemental services [9]
toddler [11]
validity [55]

Classification of Day Care Services

Day Care [1] is defined as care provided to a child by a person or persons outside the child's immediate family, either inside or outside the child's home.

- A day care center [2] is defined as a licensed facility in which care is provided to 13 or more children under the age of 13, generally for up to 12 hours each day, five or more days each week, on a year-round basis.

- The term family day care home [3] refers to a private family home, generally not licensed, in which children receive care, usually for up to 12 hours each day, five or more days each week, on a year-round basis. Most state licensing codes limit family day care homes to a maximum of six children.

- A group day care home [4] is defined as a private home serving 7 to 13 children, with one or two adults.

- In-home day care [5] is defined as care provided to a child in the child's own home by a nonrelative or by a relative who is not a member of the child's immediate family.

Day care of any of these types may be either full-time or part-time.

- Full-time day care [6] is defined as care for 30 or more hours per week.

- Part-time day care [7] is defined as care for less than 30 hours per week.

The services provided by a day care center may be classified into two blocks.

- Care care [8] refers to the common components of the daily experience of all children in day care centers. Core care includes provision of meals, snacks, space and educational/play materials, arrangements for minimum health care, and various caregiver services necessary to the nurturance of young children.

- Supplemental services [9] are those services to children and their families provided by a day care center in addition to core care. For children, such services include transportation, diagnostic testing and referrals. For parents, examples are social, welfare and employment services, and parent involvement in advisory and decisionmaking capacities. Supplemental services often address fundamental needs; the term "supplemental" merely reflects the fact that they are outside the scope of a minimal center day care program.

Children and Staff

The following terms are applied to children and adults in day care settings.

- Preschoolers [10] are defined as children three, four and five years of age (36-71 months). In some states most five-year-olds attend kindergarten and thus are considered school-aged children. In these cases, preschoolers are predominantly 36 through 59 months of age.

- *Toddlers* [11] are defined as children aged 18 through 35 months of age.

- *Infants* [12] are defined as children from birth through 17 months of age.

- *A caregiver* [13] is a person who provides direct care to children in a day care center classroom, a family day care home, or in a child's own home. Unless otherwise specified, the terms *caregiver* and *staff* [14] are interchangeable in NDCS documents.

- *A lead teacher* [15] (or *lead caregiver* [16]) is the principally responsible caregiver in a day care classroom. The term "teacher" is not intended to connote a school-like atmosphere in the day care center. The term *caregiver* has been used to refer to persons working with children in day care settings, and the term *lead teacher* is sometimes used to distinguish the principally responsible caregiver in a day care classroom from her aides.

- An *aide* [17] is a caregiver who assists a lead teacher in a day care classroom.

- A day care *provider* [18] is a person who is directly or indirectly involved in the provision of day care services; including caregivers, center directors and owners.

Classification of Day Care Centers

Day care centers are classified according to *legal status* [19] as profit or nonprofit.

- *Profit centers* [20] are further classified according to auspices [21] as independent centers or group centers.

  --*Independent centers* [22] are not part of a chain of day care centers.

  --*Group centers* [23] belong to a chain (group) of day care centers.

- Nonprofit centers [24] are classified according to auspices [25] as independent centers or sponsored centers.

  --Independent centers [26] are not sponsored by any group or agency.

  --Sponsored centers [27] are classified as either private or public, according to the nature of the sponsoring agency.

    --Private centers [28] are sponsored by a private agency, such as a church. (Note that all profitmaking centers, as well as independent nonprofit centers, are necessarily private.)

    --Public centers [29] are sponsored by some government agency, such as a city school system or a county welfare department.

In addition to classification by legal status and auspices, day care centers may be classified by a cross-cutting typology according to funding source. [30]

- Parent-fee centers [31] derive more than half of their income from parent fees.

- Publicly funded centers [32] derive their funding principally from government subsidies and gifts and contributions.

Alternatively, centers may be classified by funding source [33] according to federal financial participation (FFP). This typology was used in Supply Study analyses, and the reader may find these terms used when Supply Study data are referred to.

- An FFP center [34] is defined as any center which serves one or more federally subsidized child(ren).

- A non-FFP center [35] is defined as a center which serves no federally subsidized children.

NDCS Independent Variables

NDCS independent variables [36] are those variables whose costs and effects were to be measured. There are two types of independent variables: policy variables and background variables.

- Policy variables [37] are those characteristics of day care centers which may influence the quality and cost of center day care and which are or can be affected by federal policy. The NDCS was concerned with two major classes of policy variables: classroom composition and caregiver qualifications:

  --Classroom composition [38] describes configurations of caregivers and children in day care classrooms. Classroom composition is defined by three variables. (Note that any two of these variables mathematically define the third.)

    --Number of caregivers [39] is defined as the total number of caregivers assigned to each classroom. (The term staffing pattern [40] may refer not only to the number of caregivers assigned to a classroom, but also to the mix of teachers and aides or to the mix of qualifications of the caregivers in a classroom.)

    --Group size [41] is defined as the total number of children assigned to a caregiver or team of caregivers. In most cases, groups occupied individual classrooms or well-defined physical spaces within larger rooms. In a few "open classroom" centers, children were free to move from group to group. In such cases, clusters of children participating in common activities under the supervision of the same caregiver or team of caregivers were considered to be "groups." (The term activity subgroup [42], by contrast, refers to the actual number of children interacting with a particular caregiver. A group of 20 children, for instance, might be divided into three activity subgroups, one with the lead teacher, and two with aides.)

--Staff/child ratio [43] is defined as number of caregivers divided by group size. Higher, or more stringent, staff/child ratios are those with a smaller number of children per adult. For instance, a ratio of 1:5 is higher, or more stringent, that a ratio of 1:10 (which is lower, or less stringent). Note that the terms staff/child ratio and caregiver/child ratio [44] are interchangeable in NDCS discussions.

--Caregiver qualifications [45] variables were developed to describe caregivers' years of formal education, amount of training and/or education related to child development, and amount of work experience as a caregiver.

- Background variables [46] are characteristics of day care centers which can be influenced by government regulation only indirectly, if at all. Examples are age, sex and race of children, or socio-economic characteristics of families and of the community served by a center.

NDCS Dependent Variables

NDCS dependent variables [47] are those features of day care costs and quality measured as indicators of the effects of such center characteristics as group size, staff/child ratio and caregiver qualifications (the study's independent variables).

- In NDCS discussions, the term effects [48] is often used to distinguish dependent variables pertaining to quality in day care from dependent variables pertaining to day care costs. There are two major classes of effects variables.

--The term classroom process [49] (or process [50]) refers to the behavior of children and caregivers in the classroom; that is, the dynamics of their interaction. Process was recorded using two observation instruments, one concentrating on children's behaviors (the Child-Focus Instrument) and one concentrating on caregivers' behaviors (the Adult-Focus Instrument).

11

--The term child outcomes [51] (or developmental outcomes [52], or outcomes [53]) refers to children's gains in school-readiness skills; although a number of tests and ratings of social and cognitive development were field-tested, ultimately only two, both standardized cognitive tests, proved reliable enough to be used as outcome measures: the Preschool Inventory (PSI) and the Peabody Picture Vocabulary Test (PPVT).

- Cost variables [54] correspond in the main to commonly used terminology in accounting and economics. Where terms or variables peculiar to the NDCS are introduced, they are explained in the text.

Statistical Terminology

- The validity [55] of a measure is the degree to which it measures what it purports to measure. Various features of a measure may be indicative of its validity; such as: (1) a direct conceptual relationship between the measure and the construct of interest (e.g., between an observer's count of the number of children present in a class and the variable group size); or (2) agreement with other measures of the same construct (e.g., agreement between observation-based measurements of group size and schedule-based measurements of group size).

- The reliability [56] of a measure is the degree to which it gives consistent results when applied in a variety of situations; that is, the degree to which it is free of measurement error. Reliability coefficients vary from 0.00 to 1.00. A coefficient of 0.00 indicates a completely unreliable measure; a coefficient of 1.00 indicates a measure that gives perfectly consistent results across all situations. Thus, a reliability coefficient of .95 indicates that 95 percent of the measured variation among the objects of measurement (e.g., among children) is attributable to genuine differences among the objects of measurement, and that only 5 percent of the variation measured is attributable to random effects of errors of measurement.

- The generalizability of a measure [57] is a sophisticated extension of the concept of reliability in psychological measurement theory. It incorporates the notion that the numerous sources of variation in measurement groups as "measurement error" according to standard reliability theory may or may not be defined as "error," depending on one's purpose in using a given measure. [The concept of generalizability is a very complex one which cannot be clearly presented in the limited space available here. For a definitive treatment of the subject, the reader is referred to L. Cronbach, G. Gleser, H. Nanda, and N. Rajaratnam, The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles (New York: John Wiley & Sons, Inc., 1972).]

- The generalizability of a sample [58] is the degree to which the sample accurately represents a universe to which findings based on the sample are to be extended.

- The correlation [59] (degree of association) between two variables is represented by a correlation coefficient expressed as a decimal fraction. Correlation coefficients range from +1.00 (representing a perfect positive correlation) through zero (representing the absence of any correlation) to -1.00 (representing a perfect negative correlation). For example, a positive correlation between children's scores on Tests A and B would mean that children with high (or low) scores on Tests A also tend to have high (or low) scores on Test B. If the two tests' scores were negatively correlated, then high scores on Test A would tend to be associated with low scores on Test B, and vice versa.

- Regression [60] analysis is a technique for extracting from data an idealized representation, in the form of a straight line, of the relationship between two variables. That is, regression defines the particular straight line which is the "best" linear approximation of the less clearcut pattern exhibited in the data. Similarly, multiple regression [61] analysis extracts an idealized representation of the relationships between a given dependent variable and two or more independent variables.

- *Regression* [60] analysis is a technique for extracting from data an idealized representation, in the form of a straight line, of the relationship between two variables. That is, regression defines the particular straight line which is the "best" linear approximation of the less clearcut pattern exhibited in the data. Similarly, *multiple regression* [61] analysis extracts an idealized representation of the relationships between a given dependent variable and two or more independent variables.

- *Principal components analysis* [62] produced alternative weighted combinations of variables ("principal components"), thus allowing the researcher to select a small number of components which convey most of the important information in a data set--that is, which together account for a large proportion of the variance in the data. For example, a large number of variables related to socioeconomic status might be reduced to a few components--clusters of variables which are highly correlated with one another and only weakly related to variables in other components.

FOREWORD

Providing sound research which supports social policy directions affecting the lives of children and families is unquestionably a major goal of the Administration for Children, Youth and Families. By producing a clear signal in an often times cloudy environment, we are able to fulfill this important responsibility that has been entrusted to us.

The National Day Care Study (NDCS) is an outstanding example of our meeting this responsibility. This study has been widely recognized in both public and private sectors as one of the most important social policy research investigations ever by the Department. Its information has been widely used by many people and organizations, and it already has had a major impact on the drafting of the new HHS Day Care Regulations.

The NDCS searched for day care center characteristics which can both protect children from harm as well as foster their social, emotional and cognitive development. It discovered that these outcomes are clearly attainable when groups of children are small and when caregivers receive training in child-related areas. It also found that relaxing the staff/child ratio would not adversely affect children but could lower costs substantially and thus enable more children to receive care. That these findings held up across diverse sites and with different groups of children, provided support that all children can benefit from a single set of standards.

In all, I feel that the NDCS has more than justified the tremendous energy and time that has gone into it. Through this kind of commitment to excellence in its research programs, the Administration for Children, Youth and Families can be an instrumental force in enhancing the well-being of all children and families.

I am pleased to present the final volumes of the study--Volumes II and IV-A, B and C. Volume II is the research companion to Volume I--"Children at the Center." It provides quantitative support to the study's major findings. Volume IV is a compendium of technical papers which address study-related background issues, NDCS measures and methods and detailed results of individual outcome areas.

Jack Calhoun
Commissioner, Administration
for Children, Youth and Families

October, 1980

PREFACE

The federal government has become a major purchaser of child care, chiefly for the children of the working poor. With the growth of federal expenditures has come increased public concern about the quality and cost of care purchased with federal dollars. The National Day Care Study (NDCS) addressed this dual concern. Commissioned in 1974 by the Office of Child Development,* the study was conducted by two private research organizations--Abt Associates Inc. and SRI International. The study concluded that, by setting appropriate purchasing standards, the government could buy better care at lower cost than it currently buys, thus allowing it to serve more children within existing budgets.

Results of the study were summarized in a report published in March 1979.[1] The results were heavily cited in supporting arguments for proposed federal regulations, which were published in the Federal Register in early 1980.[2]

The present volume is one of a series supplementing the summary report.[3] It is intended to provide professionals in developmental psychology and related fields with a description of the methods and findings underlying the study's conclusions about links between regulatable characteristics of day care centers and the experiences and development of preschool children in center care.

Policy Context of the NDCS

Public concern with the quality of federally subsidized child care is embodied in the Federal Interagency just mentioned), there was little evidence of major heterogeneity that might suggest that the effects of group size are site-specific. Moreover, there was no clear numerical point of demarcation between small, "good" groups and large, "bad" ones. Most of the study's centers maintained groups of three- and four-year olds that varied in size from 12 to 24; typically, desirable behaviors decreased in frequency by roughly 20 percent, and undesirable behaviors increased by 20 percent, as group size increased within this range.

---

*The Office of Child Development is now the Administration for Children, Youth and Families (ACYF).

Third, staff/child ratio was also related to some aspects of interaction in the classroom, but the correlates of this critical policy variable, the focus of much of the controversy surrounding day care regulations, were less widespread than those of group size. Ratio was most clearly related to caregiver behavior: lead caregivers in high-ratio classes (those with few children per adult) showed essentially the same pattern of behavior reported above for caregivers in small groups. (However, the confounding of ratio and group size for the lead caregiver sample made it unclear whether the behavior pattern should be attributed to ratio, group size or both.) In addition, lead caregivers in high-ratio classes spent less time in overt management of children than those in low-ratio classes. They also spent more time interacting with other adults and in other activities not directly involving children. Thus some of the "contact time" potentially available to children by virtue of high adult/child ratios was spent in other ways. High ratios were not associated with high frequencies of one-to-one interaction between adults and children; in fact, ratio showed few systematic relationships to the behavior of children at all. Nor was ratio related to children's test score gains, except in a few isolated instances.

Title XX FIDCR. That report, issued in 1978, concluded that federal regulation was an appropriate means of maintaining quality in subsidized care but that the existing FIDCR were in need of revision.[5]

The Office of Child Development (now ACYF) had initiated the NDCS before the controversy over the Title XX FIDCR erupted. The NDCS and the Appropriateness Report were entirely independent efforts. Nevertheless the authors of the Appropriateness Report made heavy use of early results from the study, incorporating a preliminary report of NDCS findings [6] as an appendix to their own report. Subsequently, NDCS staff and the government project director were consulted during the drafting of revised regulations, which began within ACYF and was completed by the Office of HEW's General Counsel. The influence of the study is clearly visible in the proposed new standards regarding caregiver qualifications and group composition (group size and staff/child ratio). While the proposed standards deviate from the specific numerical recommendations regarding ratio and group size that appeared in the NDCS 1979 summary report, basic principles are retained--notably joint regulation of ratio and group size, with increased emphasis on the latter--as are many detailed suggestions regarding methods of monitoring and enforcement.

NDCS Approach and Findings: An Overview

The 1968 FIDCR were based on the advice of practitioners and experts in fields related to child care, as well as the best research evidence available at the time. However, in 1968 there existed only limited empirical evidence to support the basic but tacit assumptions that link various provisions of the regulations to quality of care--for example, the assumption that maintaining high staff/child ratios (few children per caregiver) will increase the quantity and quality of adult-child interaction. Nor were there data to support the assumption that regulatory control over such center characteristics as staff/child ratio, group size and staff qualifications would produce similar outcomes for children across the regions, states, sponsoring agencies and socioeconomic groups affected by federal legislation. Similarly, though a good deal was known about the different components of cost in day care, no specific evidence existed to link costs to regulated center characteristics or to quality. The NDCS attempted to fill these gaps in knowledge by identifying costs and effects associated with variations in center characteristics that were regulated or could potentially be regulated by the federal government.

The study's sponsors and designers recognized that national policymakers have many different views of the goals of day care. For example, federally subsidized day care can be seen primarily as an institution designed to free parents to work or to employ welfare recipients. However, ACYF has long been committed to the view that day care can and should foster the development of children. Hence the study focused on the quality of care from the point of view of the child--i.e., on the nature of the child's experience in day care and on the developmental effects of that experience, as measured by naturalistic observations and standardized tests. While many potentially regulatable center charcteristics were examined, primary attention focused on those characteristics which seemed most central to existing regulations and most likely to affect the daily experience of the child, namely staff/child ratio, group size and staff qualifications.

Perhaps the most general and important finding of the study was that variations in regulatable center characteristics do make a difference in the well-being of children. In contrast to many earlier studies of the effects of variations in curriculum or resource outlay in education, the NDCS showed clearly that it matters how day care classes are arranged and who staffs them. To be sure, much of what goes on in day care is not influenced by regulatable center characteristics. There is a great deal of variability in the quality of human interaction in day care settings even when the composition of the classroom and the qualifications of caregivers are fixed. Nevertheless regulatable characteristics show relationships to meaures of children's experience and of developmental change that are significant both statistically and substantively.

More specifically, for preschool children (ages 3-5), the smaller the group in which children are placed, the more they tend to engage in creative, verbal/intellectual and cooperative activity. Also, children in small groups make more rapid gains on certain standardized tests than do their peers in larger groups. When groups are larger, individual children tend to "get lost," i.e., to wander aimlessly and to be uninvolved in the ongoing activity of the group. These findings hold even when staff/child ratios are relatively high (i.e., when there are few children per caregiver).* Adding adults (usually teachers' aides) to a large group of children improves the adult/child ratio but does not necessarily result in increased engagement on the part of the child, nor improved test score gains. Significantly, children do not appear to experience more one-to-one interaction with adults when ratios are high than when they are low.

---

*In day care classrooms, unlike many public school classrooms, it is not usual to find a single adult in charge. Configurations of two or three caregivers, usually a teacher plus aides, are more common. Both the number of children and the number of adults varies significantly from classroom to classroom. It is for this reason that staff/child ratio and group size can vary more or less independently and must be examined separately. It cannot simply be assumed that large classes will have low ratios nor that small classes will have high ratios.

The behavior of caregivers toward children is also related to group or class size, but it is related to the staff/child ratio as well. In small classes and/or classes with high ratios (few children per caregiver), staff tend to devote their attention to small clusters of 2-7 children, rather than to large clusters of 13 or more. Staff in such classes also spend less time observing children passively than do caregivers in large classes and/or classes with low ratios. In addition, the staff/child ratio shows some relationships to caregiver behavior that are not found for group size. High ratios appear to make management of children easier. Also, in high-ratio classes adults spend more time with other adults and in activities not involving children, such as performance of routine chores. This outcome may suggest that high ratios benefit caregivers by providing contact with other adults and time to do necessary tasks, but it also suggests one reason why high ratios do not appear to affect the amount of one-to-one interaction between caregivers and children: in high-ratio classes some of the time potentially available for children is diverted to activities in which children are not directly involved.

On balance, NDCS findings suggest that the importance of group size as a regulatory device for influencing quality in child care may have been underestimated and the importance of staff/child ratio somewhat overestimated. This conclusion, of course, is not an argument for abandoning regulation of staff/child ratio. Not only did ratio show some positive effects, but the range of ratios examined in the NDCS was relatively narrow and relatively high. (Most centers in the study maintained classes with five to nine children per caregiver.) This range was chosen to illustrate effects of variations in ratio between levels required by the FIDCR and levels permitted by most states. Consequently, generalization of the findings to levels outside the range established by current regulatory variations is unwarranted. Moreover, a subsidiary study of center care for children under three suggested that ratio was as important as group size in influencing quality of care for infants and toddlers. Thus, while the findings suggest that controlling ratio alone is not an effective regulatory strategy, they also suggest that ratio should be included with group size in regulations governing classroom composition.

In addition to the above findings on group composition, the NDCS showed that qualifications of caregivers also affect quality of care. While years of formal education, degrees attained and years of experience per se made no discernible difference in quality of care, those caregivers who had education or training specifically related to young children (e.g., in early childhood education, day care, special education or child psychology) provided more social and intellectual stimulation to children in their care than did other caregivers, and the children scored higher on standardized tests.

To arrive at policy recommendations, these findings were integrated with results from other components of the study which were concerned with the costs associated with the various regulatable center characteristics and with prevailing practices in staffing and group composition among centers nationally. The costs of maintaining small groups and of employing staff trained or educated in child-related fields were found to be small, whereas the costs associated with maintaining high staff/child ratios were significant. Consequently it was recommended that, for preschoolers, the group size standards of the existing FIDCR be maintained or made more stringent, while the ratio requirements be relaxed slightly. The expected result would be an improvement in the quality of care for preschoolers together with a reduction in costs relative to those that would prevail if the Title XX FIDCR were enforced. Implementation of the NDCS recommendations would not require major disruption of current practice, since a high proportion of centers nationally already maintain both relatively small groups and staff/child ratios that are only a little less stringent than those mandated by the FIDCR,* despite claims of some providers and state Title XX administrators that the FIDCR ratios are unrealistically strict.[7] For infants and toddlers, institution of a group size standard and maintenance of the current ratio standard were recommended. It was also recommended that training or education in a child-related field be required of all individuals providing direct care to children, and that states be required to make such training available.

Organization of Technical Appendices

Technical Appendices to the National Day Care Study are divided into three volumes. Volume IV-A, Background Materials, contains three papers that help to set a context for overall study results: "Research Issues in Day Care, A Focused Review of the Literature," "Case Studies of the National Day Care Study Sites: Atlanta, Detroit and Seattle," and "The National Day Care Study from the Prospective of Black Social Scientists: Reflections on Key Research Issues." Volume IV-B Measurement and Methods provides seven papers that describe technical tasks undertaken to support the effects analyses reported in Volume IV-C. Included are papers about "Comparing Alternative Measures of Classroom Composition," "A Psychometric Analysis of the National Day Care Study Phase III Child Test Battery,"

---

*Staff/child ratios nationwide, averaging over all classes and ages of children, are 1:6.8, compared to 1:6.3 required by the FIDCR, and 1:12.5 permitted by state licensing requirements.[8]

"Investigation of Teacher Rating Scales Considered for Use in the National Day Care Study," "An Analysis of the CDA Checklist Data," "Interviews with Parents," "The Classroom Environment Study," and "The Econometric Model."

Volume IV-C, Effects Analyses, presents the results of analyses that investigated relationships between policy variables, classroom processes and child outcomes. Six papers are included: "The Adult-Focus Observation Effects Analysis," "The Child-Focus Observation Effects Analysis," "Analysis of Test Score Growth in the National Day Care Study," "Classroom Process-Child Outcome Analyses," "The Atlanta Public Schools Day Care Experiment," and "The Effects of Day Care in Eight Atlanta Public Schools Day Care Centers." All of the papers in the Technical Appendices were prepared by study analysts and were the basis for findings presented in Volumes I and II.

PREFACE

1. Ruopp, R., Travers, J., Glantz, F., and Coelen, C. Children at the Center. Final Report of the National Day Care Study: Summary Findings and their Implications. Cambridge, Mass.: Abt Books, 1979.

2. Federal Register, March 19, 1980.

3. Other supporting volumes include Coelen, C., Glantz, F., and Calore, D. Day Care Centers in the U.S.: A National Profile 1976-1977. Cambridge, MA: Abt Books, 1978; and three volumes of Technical Appendices to the National Day Care Study. Cambridge, Mass.: Abt Associates Inc., 1980.

4. Ruopp, et al., op. cit., Appendix A. p. 231.

5. Assistant Secretary for Planning and Evaluation, Department of Health, Education and Welfare. The Appropriateness of the Federal Interagency Day Care Requirements: Report of Findings and Recommendations. Washington, D.C.: U.S. Government Printing Office, 1978.

6. Travers, J., and Ruopp, R. National Day Care Study Preliminary Findings and Their Implications Cambridge, Mass.: Abt Associates Inc., 1978.

7. See Ruopp, et al., op. cit., Chapter 8, 155, and Appendix A, 230-240.

8. Ruopp, et al., op. cit., Appendix A., 236.

ACKNOWLEDGEMENTS

Since the National Day Care Study (NDCS) began in 1974, a great number of people have participated in the effort. These include project staff at Abt Associates Inc., site staff in Atlanta, Detroit and Seattle, and a panel of consultants from across the country, all of whom were ably directed by Dr. Richard Ruopp, Project Director. Staff and consultants at the Administration for Children, Youth and Families, and in particular Mr. Allen Smith, the Government Project Officer, also provided valuable direction for the study. Individual staff and roles are acknowledged in greater detail in Volume I, Children at the Center.

The final task for the NDCS has been the preparation of these volumes of Technical Appendices. The authors of the papers contained herein wrote, rewrote and revised their individual papers to prepare them for publication. In many instances Dr. Lorelei Brush and Dr. Jeffrey Travers gave special technical direction and Sally Weiss provided editorial assistance. Nonetheless, each paper represents an individual effort by each author. No attempt has been made to ensure consistency of style or format or to link the findings of the various papers; this was the purpose of Volume I, Children at the Center and Volume II, Research Results.

Producing all of these papers has required a considerable effort by Karen Hudson, secretary for this final phase of the study, and Christine Bornas, secretary during the earlier phases. They have managed to prepare drafts, organize changes, make corrections and produce the final papers, always within the time schedules provided. To them, and to all of the authors and support staff, I give my warmest thanks.

Nancy Goodrich
Technical Coordinator
Abt Associates Inc.
Cambridge, Mass.
October, 1980

The Adult-Focus Observation Effects Analysis

Barbara Dillon Goodson

TABLE OF CONTENTS

Page

CHAPTER ONE: INTRODUCTION TO THE STUDY 4

Measuring Caregiver Behavior 5
Study Design 7
Summary 12

CHAPTER TWO: OBSERVATIONS OF CAREGIVERS 14

Phase III Samples and Procedures 14
The Adult-Focus Instrument 16
Introduction to the AFI Analyses 18

CHAPTER THREE: THE CAREGIVER IN THE CLASSROOM: DESCRIPTION OF CAREGIVER BEHAVIOR 22

Content of Caregiver Interactions 22
Object of Caregivers' Interactions 26
How Caregivers Interacted 28

CHAPTER FOUR: ANALYSIS OF CAREGIVER OBSERVATION DATA: DEPENDENT AND INDEPENDENT MEASURES 31

Selection and Construction of Dependent Measures 31
Reliability of the Dependent Measures 32
Selection of Independent Measures 41

CHAPTER FIVE: EFFECTS OF THE POLICY VARIABLES 46

Experimental Study of Ratio Effects 46
Regression Model Construction 48
Regression Results: Lead Teacher Behavior in the 49-Center Study 56
Regression Results: Aides in the 49-Center Study 81
Discussion of Findings for the 49-Center Study 101

TABLE OF CONTENTS (cont.)

Page

Regression Results: Fall AFI Analyses in the 49-Center Study 104
Regression Results: Caregivers in the Atlanta Public Schools (APS) Centers 108

References 109

CHAPTER ONE: INTRODUCTION TO THE STUDY

The overall objective of the National Day Care Study (NDCS) was to provide empirical information in three areas of major policy concern: (1) the quality of day care provided, (2) the per-child cost of the day care provided; and (3) the potential quality/cost trade-offs associated with alternative formulations of federal day care regulations. The study focused on the largest group of children receiving federally subsidized care--preschool children (aged 3-5)--and on the day care settings in which most of these children are found--urban day care centers serving low-income families. The study also focused on program characteristics that have long been considered key determinants of quality and cost in center care--staff/child ratio, group size and caregiver qualifications. The mandate of the NDCS was to examine the impact of variations in these and other regulatable characteristics on the quality and cost of care received by preschool children.

The NDCS addressed the following major policy questions:

- How is the daily experience and consequent development of preschool children in day care centers affected by variations in staff/child ratio, group size, caregiver qualifications and other regulatable center characteristics?
- How is the per-child cost of center-based day care affected by variations in staff/child ratio, group size, caregiver qualifications and other regulatable center characteristics?
- How does the cost-effectiveness of center-based day care change when adjustments are made in staff/child ratio, group size, caregiver qualifications and other regulatable center characteristics?

This paper presents findings relevant to the first-policy question, specifically, how is caregiver behavior in day care centers affected by variations in the key program characteristics (group size, ratio, and caregiver qualifications).

Measuring Caregiver Behavior

Caregiver behavior was of crucial interest in studying quality of day care since, for the children in a center, the caregiver is a primary determinant of their experience. The day-to-day interaction between caregivers and children, and the caregiver's methods of organizing and managing the program, are the central mechanisms by which the center environment influences children's learning and development. The aspects of day care centers that are key in the NDCS--group size, ratio, and caregiver qualifications--are presumed to be linked to quality of care precisely because they affect the nature of caregiver/child interactions.

Although it was presumed that caregiver behavior was an important aspect of quality in day care, there was no established theory specifying which caregiver behaviors promote child development in this setting. At the same time, there exists a wealth of research findings which have at least potential application to study of caregiver behavior in day care settings. Selection of the types of caregiver behavior to be studied in the NDCS was based, in part, on data from four broad areas of research: studies of how caregiver behavior is related to center characteristics; research on adult (particularly parent) behavior which promotes child development; research on teacher effectiveness with children in early grade school; and descriptions of day care environments. The existing studies, largely based on settings other than day care, point to the importance of certain general dimensions of adult behavior that may be relevant for day care as well. For example, investigations of elementary school teaching suggest that "warmth" and "planfulness" are related to student effort, achievement, creativity and interest in school. Studies of parent-child interaction suggest that development of intellectual skills is associated with a range of parental behaviors, socialization practices, and aspects of the emotional climate of the home. Included among these are frequent conversational interchange, especially contingent verbal exchanges, reading to the child, and encouragement of the child's autonomy and participation in decision-making.

The available research thus suggested both macro-variables describing patterns of interaction (proportion of time interacting with various size groups) as well as qualities of interaction (e.g., warmth), and micro-variables such as contingent verbal response, and rational explanations. All of these macro- and micro-variables could, and perhaps should, be measured in a study of quality of day care.

On the other hand, the NDCS operated in actual day care settings with substantial variety in their environments, and one of the study objectives was to describe links that generally had not been studied previously. Therefore it seemed wisest to try to obtain a broad-brush picture of caregiver behavior in actual day care settings, focusing on broad patterns of interactions assumed to be especially sensitive to classroom composition variables, and on broad qualities, such as amount of direct interaction with children versus more passive supervision. To that end, the observations were focused on describing the daily process in the day care classrooms, including what caregivers did with children (e.g., teach vs. observe), their verbal techniques (e.g., command vs. question), and whom they worked with (one child vs. larger groups).

Direct observation of caregivers in day care center classrooms was the major method used to measure caregiver behavior. Use of observations to study behavior in natural settings such as day care is a procedure that has strong intuitive appeal. The connection between data and phenomena is unusually direct. Natural observations avoid the artificiality that opens many laboratory studies to the charge that their findings have nothing to do with real-world behavior. Use of such observations in the NDCS exemplifies the "ecological" approach to the study of child development urged by some of the field's most prominent spokesmen, notably Urie Bronfenbrenner.

Study Design

The NDCS was conducted in 57 day care centers in three sites.* Atlanta, Detroit and Seattle were chosen as the study sites; and a total of 57 centers subsequently were selected for participation. The centers were selected for high or low values of staff/child ratio, group size and staff education.

The study covered three years. Site and center selection were carried out during Phase I of the study, July 1974 to September 1975. The year following site selection was devoted to a study of naturally existing relationships between regulatable center characteristics and measures of classroom process, including caregiver observations. This phase of the study was used to (1) formulate initial hypotheses about relationships among regulatable center characteristics, classroom process and developmental outcomes; and

---

*See NDCS Phase I Final Report (1976).

(2) refine the measures of regulatable characteristics, classroom process and developmental outcomes to be used in in the third and final year of the study.*

Phase III (October 1976 to September 1977) was designed to answer the study's three major policy questions. The Phase III investigation had two components: a quasi-experiment conducted in 49 of the centers in all three sites, and a randomized experiment conducted in eight centers operated by the Atlanta Public Schools (APS). These eight APS centers were not included in the 49-center sample. In both studies, selected center characteristics were altered systematically, permitting measurement of the effects on caregiver behavior associated with such changes. Also, there was sufficient natural variation in all of the policy measures to allow examination of the associations between the unaltered measures and caregiver behavior.

49-Center Quasi-Experimental Study

The quasi-experiment in 49 of the centers was a comparison of three groups of centers (Table 1.1). Group I (the "treatment" group) consisted of 14 centers which had low observed staff/child ratios (1:9.1) in Phase II, and whose ratios were increased to 1:5.9 in Phase III.** Effects of this treatment on caregivers and children were compared with results from a matched group of 14 untreated centers with naturally low-ratio (1:9.1), labeled Group II, and a group of 21 untreated centers with naturally high ratio

---

* See NDCS Second Annual Report, 1975-1976.

**Note that, in conformance with HEW directives, manipulations consisted only of making low ratios higher. The Group I treatment simulates one potential effect of full enforcement of FIDCR under Title XX--namely an increase in ratios in centers serving publicly funded children but currently operating below FIDCR ratios.

Table 1.1

DESIGN OF THE 49-CENTER QUASI-EXPERIMNENT

Group I -	Treated centers (Observed mean ratio for 14 centers = 1:9.1 in Phase II; ratio raised to 1:5.9 in Phase III)
Group II -	Untreated low-ratio centers (Observed mean ratio for 14 centers = 1:9.1)
Group III -	Untreated high-ratio centers (Observed mean ratio for 21 centers = 1:5.9)

(1:5.9), labeled Group III. The question of central interest in the quasi-experiment was whether the experimentally induced increase in staff/child ratio would produce different caregiver behaviors than were seen in the matched low-ratio centers (Group I vs. Group II). A supplementary question was whether the treated and untreated high-ratio centers (Group I and II) looked different. That is, would the experimental increase in ratio eliminate most or all differences between centers that previously operated at different ratios, or would differences in outcomes continue to exist, presumably because of other center characteristics that normally accompanied high ratios but were unaffected by the experimental increase in ratio?

Ratio was chosen for manipulation because of its critical policy relevance. Group size, caregiver experience and years of education were distributed as evenly as possible across the three experimental groups, so that the effect of ratio could be clearly separated out. No attempt was made in the quasi-experiments to alter natural variations in age-grouping. The three sets of ratios applied to classrooms that served primarily three- and four-year-old children. In some centers, three-year-olds were clearly separate from four-year-olds; in others, the two ages were mixed in the same classroom.

APS Study

The APS Study was an eight-center, 29-classroom experiment in which children were randomly assigned, within centers, to classrooms that differed systematically in level of staff education and staff/child ratio (Table 1.2). Group size and caregiver experience were distributed as evenly as possible across the three experimental groups. Twelve of the experimental classrooms served three-year old

Table 1.2

DESIGN OF THE ATLANTA PUBLIC SCHOOLS (APS) EIGHT-CENTER EXPERIMENT

	High Ratio (Observed Mean Ratio = 1:5.4)	Low Ratio (Observed Mean Ratio = 1:7.4)
High Staff Education*	4 classrooms	4 classrooms
Medium Staff Education*	7 classrooms	4 classrooms
Low Staff Education*	6 classrooms	4 classrooms

*Level of education was defined by the education of the lead teacher in the classroom. High = Master's Degree; medium = a degree from Atlanta Area Tech (2-year program); low = high school degree.

children and 17 served four-year olds. This design made possible a relatively clearcut assessment of the effects and interactions of staff education and staff/child ratio for children of different ages (three- and four-year olds).

Staff in the APS centers fell into three distinct categories of educational background. First, center directors (who were required to work in classrooms as well as to function as directors) had bachelor's degrees; most also had master's degrees. Second, lead teachers were graduates of the Atlanta Area Technical School (AAT) two-year post-secondary training program in day care or had completed at least two years of college. Third, aides generally had high school diplomas (or an equivalent such as the G.E.D.); the majority of aides had also completed the 60-hour state-required training courses in day care offered through AAT. As shown in Table 1.2, persons at these three levels of education were assigned to be lead teachers in the experimental APS classrooms--some in classes with relatively high staff/child ratios, others in classes with lower ratios. Thus, ratio and education were crossed in a two-way factorial design. Children were then randomly assigned within centers to these experimentally organized classes. Random assignment, together with the fact that the children served by APS centers were unusually homogeneous in ethnic and socioeconomic background (virtually all were black children from low-income families) minimized any confounding of center characteristics and children's background characteristics.

Summary

The two Phase III components addressed the same questions but had designs with different experimental strengths and weaknesses. Because the 49-center study included a large and diverse group of centers in three different sites, its results, if uniform across the sample, were likely to be widely generalizable; however, the diversity of the sample also posed challenges for analysis and interpretation. The APS study provided a greater degree of experimental control and afforded more safeguards against confounding of center characteristics with characteristics of the children, families or communities served. However, the generalizability of its results was potentially limited by the homogeneity of the sample. The consistent results actually obtained from the two study components constitute a far sounder basis for policy conclusions than would findings from either component alone.

CHAPTER TWO: OBSERVATIONS OF CAREGIVERS

Once the decision was made to measure caregiver behavior through direct observation, it was necessary to designate which caregivers were to be observed and when, in order to obtain a reliable and representative picture of behavior, and to select an instrument appropriate to the NDCS purposes. This Chapter describes the data collection procedures and the observation system that was used to record caregiver behavior. This discussion is followed by a broad outline of the analyses of the caregiver observations.

Phase III Samples and Procedures

Observations were conducted in all 57 NDCS study centers at two times during Phase III of the study: October 1976 and April 1977. Caregivers were observed in all classrooms that enrolled a majority of three- and four-year-old children. Two hundred and ten caregivers were observed in the fall; 220 were observed in the spring.

The staff observed included both lead teachers and aides. The sample represents a total census of lead teachers in these classrooms, since typically there was only one lead teacher in a given classroom. It represents only a partial sampling of aides, however, for two reasons. First, in classrooms with multiple caregivers, only one aide was observed per classroom, regardless of the total number of aides assigned to the classroom. In the NDCS target classrooms, the average number of aides (full-and part time) was 2.8, so the total number of aides observed was a relatively small proportion (between one quarter and one third) of the total number of aides in the classrooms. Furthermore, in a number of classrooms that had multiple caregivers, only the lead teacher was observed. Second, in both the fall and the spring, approximately twice as many lead teachers as aides were observed. This discrepancy occurred primarily because aides were much less stable in attendance in the classrooms. Most worked part-time, and absence was much more frequent among aides. It consequently was more difficult to complete observations on aides. For these reasons, results of analyses of the aide sample are treated more tentatively than results for lead teachers.

In classrooms where only a lead teacher was observed, the teacher was observed for two mornings in a week. Where a teacher and an aide in a classroom were observed, each was observed the equivalent of a morning, usually on two days during a week. Observations of caregivers were restricted to the hours between 9 a.m. and noon, since this is the most stable* period of the day in terms of child and caregiver attendance. It is also the period most linked with planned educational activities, which increased the opportunities to see caregivers interacting with the children.

In the fall, all observations of an individual classroom were conducted by the same observer. In the spring, however, two observers--one white and one black--were assigned to each classroom, and the focus caregiver was observed an equal amount of time by each observer. This change in procedure permitted examination of coding differences that could be attributed to an observer's race and distributed any coding differences across caregivers and classrooms.

Twenty-one observers collected data on caregivers. Observers were selected from the local community at each site and trained by SRI International for approximately one week just before each data collection period. (A detailed description of the training may be found in SRI's Phase III Report.) At both data collection points, observers were essentially comparable on all their background characteristics except race. Most observers were female, and college graduates or soon to be college graduates; the average age was about 33 years, with observers in Detroit tending to be slightly older than the others. The primary difference between the observers hired in the fall and those hired in the spring was their race. In the fall, most observers (70%) were white, while in the spring, the number of black and white observers was almost equal in order to accommodate biracial observation teams.

---

*See Bache (1980) and Abt NDCS Phase II Research Report (1977).

The Adult-Focus Instrument

The instrument chosen to record behavior was the SRI Preschool Observation Instrument, or Adult-Focus Instrument (AFI). The AFI had previously been used by SRI International in evaluating the Follow Through and Head Start Planned Variation projects. It was modified (and hence renamed) for the NDCS to record adult behavior in day care centers. The AFI is designed to describe the day care classroom environment and to record the behavior of individual caregivers. The instrument has three sections:

- Physical Environment Inventory--a description of the equipment present in a classroom;
- Classroom Snapshot--a recording of the numbers of staff and children present at a specific point in time, and their activiites and groupings; and
- Five-Minute Interaction (FMI)--a record of the behavior of a single focus caregiver during a five-minute period.

Descriptive data from the Physical Environment Inventory were combined statistically into a single rating of physical quality for each center, and data taken from center space plans were used to confirm that NDCS centers met or exceeded state licensing requirements for indoor and outdoor square footage per child. Classroom Snapshot data were used mainly to provide group size and staff counts for computing the classroom composition measures, while the Five-Minute Interactions (FMI's) provided the bulk of the data used in the major analyses of caregiver behavior. It is through a detailed analysis of the FMI data in conjunction with the policy variables that the relationships between regulatable center characteristics and caregiver/child interaction were assessed.

The FMI's were designed to record a narrative description of caregiver behavior. Each FMI consists of five minutes of observation, broken into 63 interaction frames. Observers were allowed to set their own rate of coding on the FMI's; a maximum of 63 frames could be coded during five minutes of observation, but no minimum was set. Since each FMI covered only five minutes of observation, in an observation period multiple FMI's were completed on a caregiver.*

Each frame in the FMI is, in effect, a sentence about an action observed. It describes the actor (WHO), the object of the action (TO WHOM), the content of the action (WHAT), and the style (HOW). (In the NDCS observations, the focus caregiver being observed was either the WHO or the TO WHOM of the action in each frame of the FMI.) In each frame of an FMI, one code for WHO, TO WHOM, and WHAT had to be recorded. As shown in Table 2.1, there were 12 WHAT codes to choose from to indicate the action or behavior that was occurring. These codes are the most important in the analyses, so brief definitions are provided in Table 2.2. The 12 WHO/TO WHOM codes listed in Table 2.1 are basically self-explanatory. The HOW codes provide information about the action that is occurring, describing its content or affect. HOW codes were optional; although none or from one to six codes could be recorded per frame, the average number recorded per frame was one. The relative frequencies of occurrence of the AFI codes are presented and discussed in Chapter Three.

---

*In the NDCS study, the average length of observation of a caregiver in a single caregiver classroom was 3 hours (approximately 36 FMI's); in a two-caregiver classroom, it was 1 1/2 hours (or 18 FMI's).

Introduction to the AFI Analyses

The AFI analyses fall into three groups. First, the frequencies and variabilities of all the codes from the FMI were examined, which produced a description of caregiver behavior in these day care centers. This description helped set a context for the subsequent analyses linking caregiver behavior and the key program characteristics. Two types of linking (or effects) analyses were carried out. First, for the 49-center quasi-experiment, a series of one-way analyses of variance were done using the three groups of centers (Table 1.1) as levels of an independent, classifactory variable, and the FMI codes as dependent variables. Second, there was, within each of the various experimental groups of centers and classes, a great deal of variation not only in the experimentally manipulated variables (ratio and staff education) but also in other regulatable characteristics--group size, staff experience and content of caregivers' education or training. These naturally occurring variations were examined, though multiple regression analysis, in relation to the caregiver variables. These regressions are the central effects analyses for caregiver behavior.

Table 2.1

PHASE III AFI CODES USED IN THE FIVE MINUTE INTERACTIONS

WHO/TO WHOM

Teacher
Aide
Parent
Volunteer/Visitor
Child
Different Child
Toddler
Infant
Small Group (2-7)
Medium Group (8-12
Large Group (13+)
Other

WHAT

Commands
Direct Questions
Open-Ended Questions
Responds
Instructs
Adult Self-Related Activity or Conversation
Center-Related Statements and Activity
Supports/Comforts
Praises/Acknowledges
Corrects
No Response
Rejects
Observes/Attends

HOW

Touch
Nonverbal
Movement

Task
Behavior
Utilitarian
Negative
Happy
Guide
Punish
Sad
Dramatic Play
Materials
Rule

Table 2.2

DEFINITIONS OF "WHAT" CODES FROM THE AFI

COMMAND:	An order that asks for a response free of argument.
QUESTION:	Request for direct recall of material, a statement of preference.
RESPONSE:	Compliant response to a command, question, correction, or to praise.
INSTRUCT:	Demonstrations of activities explanation of rules, provision of information.
ADULT-RELATED ACTIVITY:	Verbal and nonverbal activity between adults that is non-center and non-child focused.
CENTER-RELATED ACTIVITY:	Statements or activities that involve children or tasks in the center.
COMFORT:	Statements or activities of affectionate attention and comfort.
PRAISE:	Approval, praise, acknowledgement, recognition, verbal or nonverbal.
CORRECT:	Attempts to change or modify a response, feeling, product or behavior.
NO RESPONSE:	A compliant response is expected but does not occur.
REJECT:	Negative, noncompliant responses, verbal or nonverbal.
OBSERVE:	Adult listens to or observes others.

An important decision made prior to any of the analyses was the choice of an appropriate level of analysis --caregiver-level or classroom-level. Since a teacher and an aide were observed in many classrooms, the observation data could be combined to form classroom-level measures; alternatively, the groups of teachers and aides could be examined separately. The latter approach was taken. The decision to do analyses at the caregiver-level was made primarily because, as previously described, the aide sample was incomplete. Because some classrooms with aides had no aide data and many classrooms with multiple aides had data from only one aide, it did not seem valid to combine the data of teacher(s) and aide(s) from the same classroom.

CHAPTER THREE: THE CAREGIVER IN THE CLASSROOM: DESCRIPTION OF CAREGIVER BEHAVIOR

The observation data provide a picture of the daily process in NDCS centers. This picture helps set the stage for the effects analyses by indicating which kinds of behavior were frequent and which were rare, and how the type of activity and the configuration of actors (numbers of adults and children) were related. The FMI data shown in Table 3.1 provide a picture of the content or quality of the interactions between caregivers and children through the WHAT and HOW codes. The TO WHOM codes describe the pattern of activity subgroups, (i.e., the number of children actually interacting with a particular caregiver, as opposed to the total number of children under supervision of the caregiver or team of caregivers).

Content of Caregiver Interactions

In terms of qualitative differences in caregiver/child interactions, the FMI WHAT codes can be organized into four broad dimensions: 1) social interaction, involving positive caregiver/child interactions, both directive and nondirective; 2) management, involving caregiver/child interactions focused on amending children's behavior; 3) observation/supervision, when the caregiver stands back and watches children; and, 4) center- or adult-related behavior, involving caregiver actions in which children are not focal. The first two dimensions represent active engagement with children; together they comprised an average of 45 percent of a caregiver's time. The latter two dimensions represent non-interactive behavior and occupy, on the average, over half of a caregiver's time (Table 3.1). In particular, an average of nearly 20 percent of a caregiver's time was spent observing/attending children, and the remaining 35 percent was spent in either center- or adult-related activity.

Table 3.1

CAREGIVERS' ACTIONS TOWARD DIFFERENT RECIPIENTS:
MEAN PROPORTIONS (STANDARD DEVIATIONS) OF WHAT AND TO WHOM CODES, SPRING AFI

(N=220)

	TO WHOM Code										
WHAT Code	1 child		(2-7) Small Group		(8-12) Medium Group		(13+) Large Group		All Children		Staff and Environment
	$\bar{x}$	(s.d.)	$\bar{x}$	(s.d.)	$\bar{x}$	(s.d.)	$\bar{x}$	(s.d.)	$\bar{x}$	(s.d.)	$\bar{x}$
Commands	.057	(.028)	.009	(.008)	.008	(.014)	.009	(.014)	.083	(.038)	
Corrects	.052	(.028)	.006	(.005)	.003	(.008)	.003	(.006)	.064	(.035)	
Instructs	.022	(.024)	.016	(.027)	.022	(.007)	.021	(.039)	.081	(.060)	
Questions	.044	(.028)	.005	(.008)	.005	(.011)	.004	(.009)	.058	(.034)	.003
Response	.016	(.013)	.000		.000		.000		.016	(.013)	.003
Comforts	.012	(.014)	.000		.000		.000		.012	(.014)	
Praises	.038	(.026)	.002	(.003)	.002	(.005)	.002	(.005)	.043	(.029)	
Center-related	.058	(.042)	.010	(.014)	.007	(.013)	.008	(.020)	.084	(.054)	.289
Adult-related	.000		.000		.000		.000		.000		.050
Observes	.024	(.027)	.048	(.055)	.046	(.062)	.085	(.112)	.203	(.114)	.007
TOTAL	.323	(.125)	.096	(.077)	.093	(.110)	.132	(.141)	.644	(.145)	.352

The social interaction and management dimensions included the codes COMMANDS, CORRECTS, INSTRUCTS, RESPONDS, PRAISES, COMFORTS, and QUESTIONS. All of these codes were positively correlated with each other, and negatively correlated with the codes representing passive caregiver behavior with children--OBSERVE AND ADULT ACTIVITY (Table 3.2). Of particular interest among the interactive behaviors was the amount of instructing which occupied eight percent of the caregiver's time. An additional 15 percent of the caregiver observations were coded as COMMANDS or CORRECTS, which represent efforts to alter behavior, manage or control children. These two codes were strongly correlated (Table 3.2). The remaining 13 percent of caregiver time was spent "warmly" interacting with children--praising, comforting, asking questions of and responding to children, a set of codes that also were highly intercorrelated. (Note that the codes COMFORTS and RESPONDS were particularly infrequent.)

The 20 percent of a caregiver's time spent observing/attending children was approximately twice as much as any other single caregiver activity with children. Observing is not inherently passive, but the pattern of correlations among the WHAT codes suggest that, within the range of frequencies observed in the NDCS, more observing meant less of almost all other activities with children. OBSERVES was negatively correlated with all of the other codes except ADULT ACTIVITY. Although intelligent observation of children is a hard-won skill of the trained caregiver, the instrument could not distinguish different types of observing by caregivers.

An average of a third of a caregiver's time was spent in activities that did not involve verbal interaction with children. Most of this time was in CENTER-RELATED ACTIVITY, such as preparing or passing out materials. Only

Table 3.2

INTERCORRELATIONS OF WHAT AND TO WHOM CODES, Spring AF1
(N=220)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
COMMAND		.41**	.30**		.20**	-.29**	-.37**		.27**	-.19**	.46**	-.09	.17*		-.17*	-.45**
CORRECT			.20**	.16**	.14*	-.23**	-.30**		.16*	-.22**	.43**	.08	.11	-.13*		-.34**
QUESTION				.22**	.29**	-.27**	-.37**	.09	.50**	-.26**	.61**	.11	.12	-.19**	-.17*	-.43**
RESPOND					.17**	-.17*	-.09	.19**	.43**	- 28**	.47**	.10**	-.07	-.22**	.11	-.23**
INSTRUCT						-.16*	-.33**	-.08	.27**	-.37**	.19**		.23**		.19**	-.27**
ADULT ACTIVITY									-.32**		-.39**	-.27**			.26**	.26**
CENTER ACTIVITY									-.31**	-.42**	-.32**		-.27**	-.27**	.17*	.78**
COMFORT									.14*	-.23**	.34**	.08	-.08	-.20**		-.09
PRAISE										-.28**	.56**	.12		-.17*	-.08	-.42**
OBSERVE											-.29**			.48**	-.10	-.27**
TO 1 CHILD												.29**		-.38**	-.08	-.60**
TO SMALL GROUP													-.25**	-.37**		-.20**
TO MEDIUM GROUP														-.38**	-.12	-.21**
TO LARGE GROUP															-.20**	-.12
TO STAFF																

Note: Correlations reported are p<.15. Correlations with one asterisk are significant at p<.05; those with two asterisks are significant at p<.01.

five percent of a caregiver's time, on the average, was spent in activities that were neither child-related nor center-related.

The pattern of caregiver behavior that emerged was strikingly similar in quality and quantity for both the fall and spring observations. It is apparent that caregivers must display some amount of management behavior, social interaction, passive observation and non-child related activities to run a day care classroom effectively.

Teachers and aides behaved somewhat differently in the classroom. Compared to teachers, aides did less commanding and instructing, and more observing (Table 3.3). This pattern is predictable, since aides in the NDCS classrooms typically acted as assistants with less responsibility than the lead teacher.

Object of Caregivers' Interactions

The broad picture of caregiver behavior gained from the WHAT codes can be expanded by studying the recipients of the caregivers' attention. Approximately one-third of caregivers' time was spent with individual children, one-third with groups of children, and the remaining third either with other staff or in solitary preparation of materials (Table 3.1). Of the time with children, about half was spent with individual children, while the remaining half was almost equally divided among small, medium and large groups of children. Teachers and aides looked very similar in how their attention was distributed (Table 3.3).

What caregivers did and whom they worked with were strongly related. The joint distribution of WHAT and TO WHOM codes suggests that different kinds of activities

Table 3.3

MEAN FREQUENCIES OF WHAT AND TO WHOM CODES AS A FUNCTION OF CAREGIVER JOB, SPRING AFI
(N=173*)

	Teachers (n=115)	Aides (n=58)	p**
WHAT Codes			
COMMANDS	.086	.070	.01
QUESTIONS	.061	.059	
RESPONDS	.022	.019	
INSTRUCTS	.090	.068	.01
ADULT-RELATED ACTIVITY	.060	.033	
CENTER-RELATED ACTIVITY	.380	.380	
COMFORTS	.015	.013	
PRAISES	.047	.047	
CORRECTS	.066	.064	
OBSERVES	.172	.240	.00
TO WHOM Codes			
TO 1 CHILD	.341	.341	
TO SMALL GROUP	.082	.117	.05
TO MEDIUM GROUP	.091	.077	
TO LARGE GROUP	.109	.118	
TO CHILDREN	.634	.653	
TO STAFF	.057	.064	
TO ENVIRONMENT	.278	.266	

*Caregivers from the Atlanta Public School centers were not included because of manipulations of job functions made as part of NDCS.

**p values recorded only if p<.15.

occurred with different numbers of children (Table 3.1). (This is also borne out in the correlations of the WHAT and TO WHOM codes shown in Table 3.2). When caregivers were instructing, they were as likely to be involved with more than one child as with individual children. Other activities--QUESTIONS, RESPONDS, COMFORTS, AND PRAISES--occurred nearly exclusively with individual children. These were "warmer" and most interactive codes. COMMANDS, CORRECTS, and CENTER-RELATED ACTIVITY occurred mostly with individual children but also with groups.

The code OBSERVES was in a class by itself; it became more frequent as the sizes of groups increased and was usually recorded between a caregiver and a large group of children. Caregivers observed/ supervised larger groups of children during free play periods; observation of smaller groups occurred in both free play and during task-oriented activities where the caregiver had structured the activity and then let the children work on their own.

Two codes--ADULT-RELATED ACTIVITY and CENTER-RELATED ACTIVITY--were used primarily to record caregivers' interactions with staff or the environment. By definition, ADULT-RELATED ACTIVITY never occurred with children. CENTER-RELATED ACTIVITY was observed with children an average of only eight percent of the time but with adults or with materials 29 percent of the time.

How Caregivers Interacted

The remaining set of FMI codes--the HOW codes--described the manner in which caregivers interacted with children. Because they were optional, all of the HOW codes were recorded infrequently, and therefore the codes were not analytically useful. Only about half of the codes had frequencies above .01 (Table 3.4). Further, the HOW codes with the highest frequencies, such as MOVEMENT, were of least substantive interest, while those most closely tied to theoretical concepts such as "warmth" were rare events.

Caregiver affect was of some interest. Overt affect--NEGATIVE or POSITIVE--was coded relatively rarely; however, POSITIVE affect was recorded more than three times as often as NEGATIVE. When the categories of POSITIVE affect and TOUCH are combined, it is clear that some positive interaction occurred in approximately eight percent of a caregiver's observations. The indicators of positive and negative affect usually accompanied direct caregiver-child interchanges. POSITIVE affect was coded most often in the context of praising. Caregivers touched children most often while comforting them. Not surprisingly, NEGATIVE affect was exhibited most often when caregivers corrected children. In fact, about 25 percent of the time that CORRECT was coded, it involved NEGATIVE affect by the caregiver; moreover, the majority of the caregivers' corrections were responses to children's behavior (or misbehavior).

56

Table 3.4

MEAN PROPORTIONS OF HOW CODES, SPRING AFI

(N=220)

	x	s.d.
Touch	.036	.03
Nonverbal	.275	.12
Movement	.181	.09
Task	.100	.08
Response to Child Behavior	.051	.03
Utility	.131	.09
Negative, Punish	.008	.01
Positive, Happy	040	.07
Guide	.008	.01
Sad	.000	
Dramatic Play	.003	.01
Materials	.028	.04
Rule	.004	.01
No Response to Child Behavior	.000	

CHAPTER FOUR: ANALYSIS OF CAREGIVER OBSERVATION DATA: DEPENDENT AND INDEPENDENT MEASURES

Selection and Construction of Dependent Measures

For the effects analyses, all of the WHAT codes were selected as dependent measures, along with those TO WHOM codes which occurred with frequency above .01. The HOW codes were rejected because of their low frequencies of occurrence and badly skewed distributions. In addition to the individual WHAT codes, two macro-codes--SOCIAL INTERACTION AND MANAGE--were constructed from the WHAT codes and used as summary dependent measures. The construction of these macro-codes is described below.

Several strategies were employed in an attempt to find patterns of caregiver behavior among the individual FMI codes that could be represented in constructs or macro-codes. The first technique used was a principal components factor analysis of the data, which revealed little underlying structure (i.e., no stable factors). The first factor derived in the factor analysis accounted for less than 15 percent of the variance; no other factor accounted for more than ten percent. The first factor presented almost exactly the same picture as the simple correlations: ADULT-RELATED ACTIVITY, CENTER-RELATED ACTIVITY, and OBSERVES had negative weights while the remaining codes had high positive weights (with the exception of COMFORTS, which had a loading of essentially zero). However, since this principal component accounted for relatively little variance, no single summary construct was formed from the FMI codes.

The lack of guidance from the factor analysis led back to the raw frequencies of the codes and their correlations, which were interpreted with the help of an empirical understanding of behavior in day care settings. Based on the correlations (discussed later in the descriptive analyses) and the conceptual relations among codes, two constructs were formed. One, labeled MANAGE, was a combination of the two highly correlated codes COMMANDS and CORRECTS. This construct represented caregiver efforts to change or control children's behavior. The second construct, SOCIAL INTERACTION, was formed by combining QUESTIONS, RESPONDS, INSTRUCTS, PRAISES and COMFORTS. The SOCIAL INTERACTION construct represents all verbal social interactions between caregivers and children, excluding managing children.

The intercorrelations of these derived constructs with the unaggregated codes, for teachers and aides combined, appear in Table 4.1. (The pattern of correlations among the constructs and the individual codes were very similar when teachers and aides were analyzed separately.) As expected, SOCIAL INTERACTION was significantly correlated with all the component codes and negatively related to OBSERVE, CENTER ACTIVITY and ADULT ACTIVITY. MANAGE also was highly correlated with its component codes and negatively correlated with CENTER and ADULT ACTIVITY. On the other hand, MANAGE was not correlated with OBSERVE or with some of the social interaction codes.

Reliability of the Dependent Measures

Effects analyses of the AFI were primarily aimed at determining to what extent variations in caregiver behavior are related to variations in group size, staff/child ratio, and caregiver qualifications. The reliability of the AFI dependent measures was important information for developing analytic strategies and interpreting the results of the analyses. Specifically, the reliability results were used in four ways: (1) they indicated the degree to which

Table 4.1

CORRELATION COEFFICIENTS OF WHAT CODES AND CONSTRUCTS, SPRING AFI*

(n=220)

	SOCIAL INTERACTION	MANAGE
Observes	-.33**	-.09
Commands	.29**	.85**
Corrects	.21**	.83**
Questions	.68**	.34**
Responds	.54**	.18**
Comforts	.19**	
Praises	.76**	.31**
Instructs	.77**	.12
Center-related Activity	-.55**	-.47**
Adult-related Activity	-.37**	-.39
MANAGE	.30**	--

*Correlations reported are significant at p<.15. One asterisk indicates p<.05 and two asterisks indicate p<.01.

observer effects contaminated the observations; (2) they indicated which individual codes were most reliable; (3) they indicated the ceilings on the proportions of variance in the AFI measures that the independent measures could be expected to explain; and (4) they helped in deciding which level of aggregation would be most appropriate for analysis.

The reliability of the AFI measures was assessed in three ways: through generalizability computations, through observer agreement (with criterion tapes and in tests of interobserver agreement), and through examination of the stability of the measures across timepoints.* In general, the reliability analyses indicated that the measures were sufficiently reliable to support the effects analyses, i.e., a significant part of the variance in the measures could be expected to be systematic and potentially explanable by the policy measures. On the other hand, the measures were not so reliable as to predict that more than moderate amounts of variance would be accounted for. The analyses also indicated that the broader dependent measures, especially the macro codes were more reliable.

Observer effects were examined by SRI International through observer agreement with criterion videotapes and a field-tested interobserver (paired) agreement. On the criterion videotapes, agreement on all AFI codes was above 70 percent. In the field test of interobserver agreement, observer pairs of one black and one white member observed caregivers in the FMI's, spaced a week apart. Rates of agreement were approximately 90 percent for WHO and TO WHOM codes. Agreement varied from 62 to 89 percent for the frequent WHAT codes that were used in effects analyses, with most of these codes in the 70-85 percent range. (The infrequent nonoptional WHAT codes generally evidenced somewhat lower levels of agreement.) HOW codes in many cases produced high percentage agreement, based on very few occurrences. Black and white observers differed in their use of certain codes; however, many of these differences were attributable to one or two observer pairs or to low overall frequencies of the codes in question. On the whole, SRI's data suggest that interobserver agreement, while far from perfect, is good enough to guarantee that recorded frequencies of AFI codes are determined mainly by factors outside the eye of the beholder.

---

*Most of the reliability tests were done on spring AFI data because in the spring, each caregiver was observed by two observers, on black and one white, which enabled clearer assessment of observer effects.

Day-to-day stabilities of code frequencies were examined for 203 caregivers who were observed on two consecutive days in spring 1977. Stability coefficients, shown in the first column of Table 4.2, are correlations between frequencies of the same code measured on successive days for the same caregivers. Modest correlations were obtained--generally around .2. These indicate some tendency for profiles of caregiver behavior to remain the same, but they also show that behavior fluctuates in response to the situation, with many caregivers showing a lot of a given kind of behavior on one day, followed by relatively little on the next day. (Low values of coefficients in Table 3.6 are to a degree artificial in that using two observers for each caregiver in the spring data collection meant changing observers from one day to the next, which contributed to to the apparent instability of codes.) However, in light of the relatively high rates of interobserver agreement obtained in SRI's field test, the relatively weak correlations shown in the table must be attributed primarily to volatility of caregiver behavior, rather than to observer differences.

Table 4.2

STABILITIES OF ADULT-FOCUS DEPENDENT MEASURES

Adult-Focus Codes/Constructs	Day-to-Day Stability (Spring 1977; N=203 caregivers)	Fall-to-spring Stability (Phase III; N=145 caregivers)
TO WHOM Codes		
To one child	.28	.26
To small group	.24	.36
To medium group	.31	.26
To large group	.40	.40
To other staff	.40	.14
To environment	.06	.24
WHAT Codes		
Commands	.13	.14
Corrects	.06	.07
Questions	.16	.27
Responds	.14	.49
Praises	.20	.47
Comforts	.19	.07
Instructs	.07	.36
Observes	.32	.38
Adult-related activity	.25	.36
Center-related activity	.06	.26
Constructs		
Management	.14	.27
Social	.22	.37

Fall-to-spring stabilities, shown in the second column of Table 4.2, were with a few exceptions approximately as high as day-to-day stabilities and in a few cases were substantially higher. (Correlations in the table are based on a sample of 145 caregivers observed in both fall and spring. Scores for each caregiver were averaged over two days of observation at each time point.) The fact that long-term stabilities do not deteriorate suggests that there is a measure of continuity in caregiver behavior as measured by the AFI. In some cases this continuity is partially obscured by short-term fluctuation.

The overall pattern of stability coefficients is mixed. Where there are low stabilities at both points, this suggests that the immediate situation controls behavior, rather than any characteristic of the caregiver. Low stabilities in the short term, together with higher long-term stabilities, suggest that there are general and long-lasting caregiver styles, but that these may be hard to detect over a short span of observation because day-to-day changes in the situation inhibit expression of the caregiver's usual dispositions. Altogether, the results of the stability analyses suggest that much of the variance in the measures of caregiver behavior may be linked to situational fluctuation and therefore may not be systematically related to the policy measures, which reflect enduring features of the day care setting. There is enough stability to warrant investigation of relationships to the policy variables; however, instability limits the strengths of the relationships that can potentially be found.

Neither interobserver agreement nor stability over time conveys all of the information needed to judge the suitability of AFI measures for use in effects analyses. What matters is how well those measures characterize caregivers, classes or centers when appropriately averaged. In

64

more technical terms, what matters is the relative importance of different sources of variance in caregiver, class and center means for each of the AFI measures. The more variance that is linked to the unit of analysis itself (caregiver, class or center), the more dependable is the measure for that unit. The more variance that is linked to other factors, including but not limited to observer disagreement and fluctuation in caregiver behavior, the less dependable is the measure as a descriptor of the chosen unit.*

To assess the relative importance of the major potential determinants of AFI scores, a variance components analysis was performed (Singer, Affholter and Goodrich, 1978). The analysis revealed that for almost all of the AFI codes and constructs, the occasion of observation was the major determinant of variation, followed (usually at a distance) by the caregiver. Center, class, observer and site all were relatively minor determinants. Thus, subject to certain qualifications*, the variance components results reinforced impressions created by findings on interobserver agreement and stability over time: Differences in observer perspective contributed relatively little to AFI scores, while fluctuations in caregiver behavior contributed heavily. This outcome is a common one in observation research (Medley and Mitzel, 1963).

---

*Note that a behavioral profile for a particular caregiver, calculated by averaging scores across different occasions of observation, is not free of distortion due to observer disagreement and behavioral fluctuation from occasion to occasion. To understand this point, consider the case in which scores recorded by two different observers are averaged to yield summary scores for a caregiver whom both have observed. (For simplicity, ignore the effects of fluctuation over time.) Averaging cancels out differences in perspective between the particular observers, each with his or her own biases, which tend to push "true" scores of caregivers up or down. The "average bias" of the particular pair of observers will in general deviate markedly from the average bias of this hypothetical universe of observers. Thus average recorded scores have (at least) two components--one linked to the caregiver, and the other to the average bias of the observer pair. By an analogous argument, scores averaged across a few occasions of observation include a component due to the average bias inherent in those occasions. Similarly, class scores produced by averaging across caregivers have a caregiver component, as well as observer and occasion components, and center scores produced by averaging across classes have class, caregiver, occasion and observer components.

Variance components results can be translated into estimates of the generalizabilities of AFI measures at various levels of aggregation (caregiver, class and center). Generalizabilities may be interpreted as proportions of variance in aggregate scores (e.g., class-average scores) that are linked to the unit in question (e.g., the class) and not to other factors. Generalizabilities of selected AFI constructs and codes are shown in Table 4.3. The table shows that AFI measures dependably characterize caregivers but are less dependable as descriptors of classes or centers. In some cases, particularly for the summary constructs, dependabilities at the caregiver level are strikingly high. The results have two obvious and important implications: first, that the caregiver is the appropriate unit of analysis for investigating effects of the policy variables, and second, that the study's ability to detect relationships between policy variables and caregiver behavior is substantial, particularly for the classroom composition variables, which are also measured with high reliability.

---

*The analysis was necessarily structured in such a way that the "occasion" component was indistinguishable from random measurement error. Also, partly because only one caregiver was observed in many classes, and partly because of certain assumptions built into the statistical procedure, it was difficult to separate variation due to classes from that attributable to caregivers. Thus the analysis tended to overstate the importance of "occasion" and understate that of "class." These biases had no substantive effect on the usefulness of the measures.

Table 4.3

GENERALIZABILITIES OF SELECTED ADULT-FOCUS MEASURES

Adult-Focus Codes/Constructs	Generalizability at Indicated Level of Aggregation		
	Center	Class	Caregiver
TO WHOM Codes			
To Small Group	.21	.23	.76
WHAT Codes			
Praises	.28	.24	.78
Observes	.23	.33	.79
Adult-Related Activity	.24	.44	.70
Constructs			
Management	.12	.34	.96
Social Interaction	.18	.44	.92

Selection of Independent Measures

Choice of independent variables was motivated by a basic value decision made at the outset of the study to focus attention on those aspects of the quality of day care that bear directly on the child. In effect, it was assumed that the primary goal of day care purchasing standards is to ensure the best possible environment for the most children. Other goals of day care--e.g., freeing parents to work, serving as a vehicle for delivery of social services to parents, employing low-income people as staff and fostering their development as professionals--were recognized as legitimate and important but were not central to the study.

As a consequence, in selecting regulatable center characteristics for intensive investigation as independent variables, priority was given to those characteristics deemed most likely to affect children's daily experiences, namely the composition of the classroom (principally group size and staff/child ratio) and the qualifications of caregivers (education and experience). Other center characteristics (space, equipment and materials; center philosophy and curriculum; director qualifications; stability of caregiver/child relationships; availability of nutrition and health services; availability of other supplementary services and specialists; opportunities for parent involvement) were examined in descriptive and exploratory fashion to determine whether any appeared to have major effects on classroom processes and child outcomes. However, in light of preliminary results which suggested that most of these variables had minimal effects on the particular outcome measures chosen, only a few of the variables were investigated further, and then only to a limited extent.

Two types of independent variables were tested in the effects analyses: background variables (such as age, sex and race of children, and socioeconomic characteristics of families and of the community served by the particular center) and policy variables (i.e., center characteristics subject to regulatory control). While background variables are unregulatable and therefore not of direct policy relevance, their effects had to be taken into account in assessing the effects of the policy variables.

Information on background characteristics of children and their families was gathered through interviews with parents. Background information included family income, sources of income, parents' education and occupation, length of parents' employment, number of siblings and number of adults living in the house. Age, sex and race of children were verified. In addition, census data were used to provide background information on demographic characteristics of the community, chiefly its socioeconomic and racial composition.

A set of variables that described the socioeconomic characteristics of the children's families were aggregated from the child level to the class level as potential independent measures. These included mean family income, mean mother's education, mean proportion of single parents, mean number of children under 12 in the home, and mean proportion of black children in the class. The simple correlations of these demographic variables with the dependent measures indicated some significant relationships, particularly with the TO WHOM codes. Since the five demographic variables were strongly intercorrelated, (Table 4.4) and because five SES covariables was a large set to enter into the regression analyses as independent measures, a clustering was completed to represent the average socioeconomic status of the children

Table 4.4

INTERCORRELATION OF SOCIOECONOMIC MEASURES IN NDCS CLASSROOMS, SPRING 1977*

	1	2	3	4	5
Mother's Education	.80	.72	-.55	-.43	-.25
Proportion of Mothers with College Degree		.58	-.33	-.24	-.31
Family Income			-.66	-.42	-.43
Proportion of Single Parents				.36	.39
Number of Children Under 12					.31
Proportion of Black Children					

*All correlations were significant at $p<.05$.

in each class. The five variables were factor analyzed and a principal component factor score was assigned to each class. This principal component, labeled CLASS SES, typically fell between -1 and 1, with positive scores indicating higher SES. The principal component was the main SES covariable in the regression equations. Caregiver race also was used as a covariable.

The major policy variables examined in the NDCS fell into two categories--those relating to classroom composition and those relating to caregiver qualifications. Three variables fell under the rubric of classroom composition:

- number of caregivers, defined as the total number of caregivers present in or assigned to a classroom;
- group size, defined as the total number of children present in or assigned to a class or to a principally responsible caregiver;* and
- staff/child ratio, defined as the number of caregivers divided by group size.

Information on caregiver qualifications was gathered from interviews with caregivers. Caregiver qualifications variables consisted of total years of formal education, presence or absence of child-related education/training, (also called "specialization) and day care experience (both prior to current job and time in current center).

---

*In all but a few NDCS centers, groups of children were assigned to particular rooms, supervised by a single caregiver or several caregivers. In a few "open classroom" centers, however, very large numbers of children (approaching 100 in extreme cases) were present in a single large room. Even in such centers, children clustered around individual caregivers or small teams dispersed around the room, though children were often free to move from group to group. Numbers of children in these smaller groups constituted the group size used for NDCS analytic purposes. Similarly, numbers of caregivers were the number of adults in physically separated groups.

Information on the classroom composition variables was gathered by two methods, one based on schedule or roster data and the other on direct observation, i.e., as part of the AFI. Schedule-based and observation-based measures of classroom composition were not always in close agreement. Differences between the two were primarily attributable to two phenomena--absenteeism and merging of classes. Because observations capture the group configurations actually experienced by the child and because they automatically take account of absenteeism and merging, observation-based measures were used in all the analyses reported in this volume.* However, because of the importance of these issues for monitoring and enforcement, comparative investigations of the two types of measures were conducted and are reported elsewhere (Bache, 1980).

---

*Three distinct sets of observation data were collected. One set was collected on a regular basis by NDCS staff employed full-time at each center during Phases II and III; this set was used in analyses of non-behavioral data. A second set of counts was made in conjunction with behavioral observations of caregivers, and a third in conjunction with observations of children; these counts were used in the corresponding behavioral analyses.

72

CHAPTER FIVE: EFFECTS OF THE POLICY VARIABLES

The discussion that follows focuses on the non-experimental study of naturally existing relationships between the policy measures and caregiver behavior, in the 49-centers. The results of the ratio experiment in these 49 centers are briefly presented first, but the essentially null effects reduce the need for elaboration. The results of the APS study are not discussed in this chapter but in a separate paper in this volume (Goodrich, 1980). In most cases, the discussion of results includes only those significant at p.05. The tables, however, report significance levels at least up to p.15. When nearly-significance results appear to be part of an overall pattern, the results are discussed as part of a *trend* or *tendency*.

Experimental Study of Ratio Effects

As described earlier, the effects of staff/child ratio were tested in a quasi-experiment in 49 of the centers. The question tested was whether caregivers in high-ratio classrooms with an experimentally induced increase in ratio behaved differently when compared with caregivers in either untreated low-ratio or untreated high-ratio classrooms. In a series of one-way analyses of variance comparing caregivers in the three types of classrooms, there were few detectable effects of ratio. These were not strong beneficial effects for higher ratios in terms of caregiver behavior.

Lead teachers and aides were examined separately. For lead teachers (Table 5.1), only one code showed a significant overall difference in frequency across the three groups. The frequency of RESPONDS was lower in naturally high-ratio centers than in the treated high-ratio and comparison low-ratio centers, which did not differ from

Table 5.1

Frequency of Lead Teacher Behavior in Three Ratio Treatment Groups

AFI Measures	Natural Low Ratio classrooms(n=32)	Treated High Ratio classrooms(n=33)	Natural High Ratio classrooms(n=42)	Significance*
WHAT codes				
MANAGE	.16	.16	.14	
Command	.09	.09	.09	
Correct	.07	.07	.05	
SOCIAL INTERACTION	.22	.24	.23	
Direct Question	.05	.06	.06	.15
Respond	.02	.02	.01	.05
Instruct	.09	.09	.09	
Comfort	.01	.01	.02	
Praise	.04	.05	.05	
Adult Activity	.07	.07	.05	.15
Center Activity	.38	.38	.40	.15
Observes	.17	.16	.18	
TO WHOM codes				
To 1 child	.33	.34	.35	
To small group	.09	.09	.09	
To med. or lg. group	.19	.21	.20	
To other staff	.06	.07	.05	.15

*p reported only if ≤.15

each other. There also was a tendency toward a lower frequency of CORRECTS in the naturally high-ratio centers. These results clearly are not attributable to the experimental manipulation, but to other characteristics of naturally high-ratio centers. For aides (Table 5.2), no differences, were significant, although one tendency potentially attributable to the ratio manipulation, appeared: Aides in treated high-ratio classrooms and naturally high-ratio classrooms devoted less attention to the physical environment than did those in low-ratio classrooms.

Regression Model Construction

The goal of the main effects analyses of the Phase III data was to define the relationships between each independent measure and caregiver behavior, and to assess how well the *set* of policy variables predicted caregiver behavior. Simple correlations alone were not adequate to test the relationships, since there was some confounding among the policy variables, e.g., ratio and group size. In addition, it was important to assess the combined effect on caregiver behavior of a set of policy variables. Data was analyzed by multiple regression, using different combinations of the policy variables that were selected to minimize the confounding within each set and at the same time, to maximize the chance of statistically separating the effects of the policy variables.

As noted earlier, ten independent measures were the focus of the AFI effects analyses. Eight were policy variables: observed group size, number of staff, ratio of staff to children, caregiver years of education, child-related education/training (specialization), previous day care experience, experience in current center, and age. Two were covariables--caregiver race, and the construct

Table 5.2

Frequency of Aide Behavior in Three Ratio Treatment Groups, Spring AFI

AFI Measures	Natural Low Ratio classrooms(n=13)	Treated High Ratio classrooms(n=20)	Natural High Ratio classrooms(n=20)	Significance*
WHAT codes				
MANAGE	.12	.13	.15	
Command	.06	.07	.08	.15
Correct	.06	.06	.06	
SOCIAL INTERACTION	.19	.24	.19	
Direct Question	.05	.07	.06	
Respond	.02	.02	.02	.15
Instruct	.07	.08	.06	
Comfort	.02	.01	.01	
Praise	.04	.06	.04	
Adult Activity	.03	.02	.04	.15
Center Activity	.42	.35	.38	.15
Observes	.24	.25	.23	
TO WHOM codes				
To 1 child	.31	.36	.35	
To small group	.09	.11	.14	
To med. or lg. group	.20	.22	.16	
To other staff	.08	.06	.06	
To environment	.06	.06	.04	.10

p reported only if ≤.15.

representing socio-economic status of the children in each class (class SES). The ten were not entered as a single group in any of the regression equations for two reasons, first, because the set was too large, relative to the sample sizes of the data sets, but more importantly because of problems of multicollinearity among the measures.

There was some substantial confounding among the independent measures, both within types of measures (e.g., the classroom composition variables) and across types (e.g., the classroom composition variables and caregiver characteristics) (Table 5.3). Three pairs of related measures were highly correlated: GROUP SIZE and RATIO, EDUCATION and SPECIALIZATION, and CAREGIVER RACE and CLASS SES. The pattern of confounding affected the strategy used in the regression analysis.

Stepwise regressions were used. However, only the final step of the regressions is reported in the tables of results, because there was no theoretical basis for predicting or interpreting the order of entry of the regressors the coefficients presented in the tables are not affected by order of entry of the independent measures.

Five regression models were tested (Table 5.4). MODEL I entered two policy variables which were not confounded: GROUP SIZE and CHILD-RELATED SPECIALIZATION. In MODEL II, RATIO was entered along with SPECIALIZATION. MODEL III entered the variables for EXPERIENCE IN CURRENT CENTER, GROUP SIZE, and SPECIALIZATION. In MODEL IV, YEARS OF EDUCATION instead of SPECIALIZATION was entered with GROUP SIZE, to compare effects for education and specialization. Finally, in MODEL V, GROUP SIZE and RATIO were entered with SPECIALIZATION to test their combined effect and to examine what happened to the separate effects of

Table 5.3

Intercorrelations of Independent Measures in the 49-Center Sample[1]

(n=87 Teachers and n=42 Aides)[2]

	Specialization	Years of Education	Previous Day Care Experience	Current Center Experience	AFI Group Size	AFI Ratio	Caregiver Race	Class SES
Specialization	--	.38**						
Years of Education	.33**	--		-.22**			.32**	
Previous Day Care Experience			--			.32**		
Current Center Experience	.23			--	-.24*	.41**		
AFI Group Size	-.25				--	-.46**		
AFI Ratio					-.65**	--		
Caregiver Race					-.27		--	.56**
Class SES				-.25			.48**	--

---

[1] All reported correlations have p≤.15. Correlations significant at p<.05 have one asterisk; those with p<.01 have two asterisks.

[2] Correlations above the diagonal are for teachers; those below the diagonal are for aides.

Table 5.4

Outline of AFI Exploratory Regressions

Model	Covariables	Policy Variables	Purpose
I	Class SES Caregiver Race	Group Size Specialization	• Estimate individual and combined effects of GROUP SIZE and SPECIALIZATION
II	Class SES Caregiver Race	Ratio Specialization	• Estimate individual and combined effects of RATIO with SPECIALIZATION
III	Class SES Caregiver Race	Group Size Specialization Experience in Current Center	• Estimate effects of caregiver EXPERIENCE
IV	Class SES Caregiver Race	Group Size Years of Education	• Estimate effects of YEARS OF EDUCATION, to compare with effects for specialization in Model I
V	Class SES Caregiver Race	Group Size Ratio Specialization	• Estimate combined effects of RATIO with GROUP SIZE

each when entered with a correlated variable. Two covariables--CAREGIVER RACE and a measure of CLASSROOM SOCIOECONOMIC STATUS (SES) were entered in every model.

The rationale for using these sets of independent measures in this manner was based primarily on the pattern of confoundings among the measures. For the three pairs of related measures which were highly intercorrelated (RATIO and GROUP SIZE, CAREGIVER SPECIALIZATION and YEARS OF EDUCATION, CAREGIVER RACE and CLASS SES), both variables in a pair could not be entered simultaneously in the regression equations if separate estimates of their effects were of interest. Therefore, a different strategy had to be adopted.

In the case of CAREGIVER RACE and SES OF THE CLASS, there was no reason to try to separate their effects, so the two measures were entered simultaneously into all regressions, and only their combined effect was examined. We assumed that there was a "package" of caregiver and child background factors that was likely to be related to caregiver behavior and that should be taken into acount. These variables were not correlated with any of the policy variables except years of education. Therefore, the effects of the policy measures could be measured accurately without entering the covariables. However, they were related to caregiver behavior and for that reason were entered for purposes of prediction.

Separate estimates of the effects for GROUP SIZE and RATIO were of interest. Therefore, in four of the five regression models, GROUP SIZE and RATIO were entered separately. In the fifth model, they were entered together. Although they were highly confounded in the AFI sample, (r = .46 for teachers, .65 for aides). GROUP SIZE and RATIO were related to somewhat different sets of caregiver behaviors. That is, the GROUP SIZE effect and the RATIO effect clearly were confounded but not synonymous, suggesting that these two policy variables should receive careful individual examination. Group size rather than ratio was focused on in most models for the sake of consistency and comparison on the basis of findings for child tests and observations, in which group size appeared to have stronger effects. The regression model testing both GROUP SIZE and RATIO showed that entering both of the confounded measures often produced artifactual effects. These included both spuriously high and reduced effects for one or both of the measures. For this reason, the results of this regression model (Model V) are not reported.

When examining CHILD-RELATED EDUCATION/TRAINING (SPECIALIZATION) and YEARS OF EDUCATION, the regression models entered the variables separately but focused on SPECIALIZATION in four of the five models. Simple correlations of these independent measures with the dependent measures were similar in direction, but those for SPECIALIZATION were stronger, particularly among lead teachers. Among aides, YEARS OF EDUCATION was about as good a predictor as SPECIALIZATION, but neither was very powerful. In addition, SPECIALIZATION showed little variance among aides. Another important reason for focusing on SPECIALIZATION was that YEARS OF EDUCATION appeared to be partially or primarily an SES measure; it was correlated with RACE OF CAREGIVER, while SPECIALIZATION was not. The regression model that tested education was used with both teachers and aides, but years of education had effects only for aides. Therefore, the results for Model IV are reported for aides only.

The caregiver experience variables were tested in models along with GROUP SIZE and SPECIALIZATION. They were not tested in models with RATIO because EXPERIENCE and RATIO were confounded in the AFI sample. For both lead teachers and aides, only the results for EXPERIENCE IN CURRENT CENTER are discussed, since PREVIOUS DAY CARE EXPERIENCE had no effects.

In addition to the conventional least-squares regression analyses, two kinds of checks were done on the AFI data to identify extreme or atypical cases which would alter the distributions required for the kinds of statistics used in the analyses. First, scatterplots of the dependent and independent measures were scanned for bivariate outliers. Second, biweighted regressions were run to assess the robustness of the regression equation if outliers were removed. In biweighted regressions, weights are assigned to cases on the basis of their deviations from the regression surface. Outliers are given less weight and thus affect the regression equation less strongly.

The scatterplots clearly indicated a handful of about seven outlier cases. These cases either had extreme values on the dependent measures (e.g., 65 percent on OBSERVE when the next highest value was below 50 percent), or extremely small values for group size. Once these cases were eliminated from the data set, the correlations between the policy and dependent measures were redone. Those that were significant before the exclusion became stronger, and some of the apparently contradictory and or unexplainable correlations disappeared. In general, however, the main effects of the policy variables were not dependent on the few atypical cases with extreme values.

Further, the estimates obtained with the biweighted regression analyses typically were not substantially different from the estimates from the unweighted regressions. Biweighted regressions were done on the spring data separately for the 49-center lead teachers. The biweighted estimates for the regressors were not much different from the unweighted estimates. Estimates for group size were virtually unchanged by the weighting, and a few of the estimates for ratio were reduced.

The regression models were investigated separately for teachers and aides, for caregivers in the 49-center and APS samples, and for fall and spring data. This report focuses on the spring data for caregivers in the 49-center sample. The focus on spring data is because the data collection techniques in the spring controlled better for observer effects and because of the instability of caregivers' classroom assignment in the fall. There also is a focus on teachers, because of the representativeness of the teacher sample. The parallel regression analyses done on the fall data and on the APS data are discussed more briefly, with emphasis on the consistency of the findings.

Regression Results: Lead Teacher Behavior in the 49-Center Study

Most of the findings for lead teachers that are reported below are based on a sample of 87 teachers. (Of the 115 teachers observed in the 49-center study, these 87 had no missing data on the variables used as independent measures in the regressions.) In the presentation of the regression analyses for this sample, the results are organized around the major independent measures. First, findings for the group composition measures (group size and ratio) are discussed for each of the dependent measures, followed by discussion of findings for caregiver qualifications, and finally, the covariables. The discussion is accompanied by tables of regression results--one table for each of the dependent measures. For each dependent measure, the table presents the findings from all three of the regression models.

Group Composition Measures

Social Interaction (Table 5.5). Larger group sizes tended to accompany fewer positive social interactions between lead teachers and children. This effect is significant ($p = .05$) in the simple regression of GROUP SIZE on SOCIAL INTERACTION, which included 115 lead caregivers.* In the multiple regression models reported here, for which $n = 87$, the relationship is no longer significant, although the direction of the relationship persists.

Among the component codes in SOCIAL INTERACTION, only RESPOND was related to group size ($p<.05$). Teachers responded to children less often in larger classrooms (Table 5.6). None of the other codes--QUESTION, INSTRUCTS, PRAISE, COMFORT--were individually associated with group size or ratio (Tables 5.7-5.10).

Management (Table 5.11). Staff/child ratio was associated with the amount of management by caregivers. Higher staff/child ratios were significantly related to less managing by lead teachers. There also was a trend for larger group sizes to accompany more managing of children by the caregiver.

Of the two component codes in Management, (Tables 5.12 and 5.13), only COMMAND was significantly related to

---

*This was reported in Volume I, Children at the Center.

Table 5.5

Results of Regressions of
Caregiver Behavior Variables, Spring 1977
(Lead Teachers, n=87)

SOCIAL INTERACTION

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.003	2.02	.17	-.14	.08
	Child-related education/ training	.052	6.55	.01	.27	(.17)
II	Observed staff/child ratio	-.110	.38	.54	-.02	.07
	Child-related education/ training	.052	6.45	.01	.27	(.16)
III	Observed group size	-.003	1.61	.21	-.14	.09
	Child-related education/ training	.055	6.66	.01	.27	(.19)
	Experience in current day care center	.000	.02	.88	.06	

Table 5.6

Results of Four Regressions of
Caregiver Behavior Variables
(Teacher-level, n=87)

Policy Variables		RESPOND Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.001	5.14	.03	-.22	.07
	Child-related education/ training	.004	3.42	.07	.14	(.32)
II	Observed staff/child ratio	.017	.587	.45	.13	.03
	Child-related education/ training	.004	2.209	.14	.14	(.28)
III	Observed group size	-.001	4.95	.03	-.22	.10
	Child-related education/ training	.005	4.50	.04	.14	(.35)
	Experience in current day care center	-.0002	.42	.52	-.05	

Table 5.7

Results of Four Regressions of Caregiver Behavior Variables
(Teacher-level, n=87)

PRAISE

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.001	.75	.39	-.09	.04
	Child-related education/training	.013	4.31	.04	.21	(.22)
II	Observed staff/child ratio	-.026	.23	.63	.00	.04
	Child-related education/training	.013	4.07	.05	.21	(.22)
III	Observed group size	-.001	1.01	.32	-.09	.07
	Child-related education/training	.014	4.56	.04	.21	(.25)
	Experience in current day care center	-.001	.38	.54	-.06	

Table 5.8

Results of Four Regressions of Caregiver Behavior Variables
(Teacher-level, n=87)

		QUESTION				
Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.0001	.02	.89	.01	.02
	Child-related education/ training	.010	2.08	.15	.19	(.10)
II	Observed staff/child ratio	.022	.134	.72	.06	.03
	Child-related education/ training	.010	2.215	.15	.19	(.11)
III	Observed group size	.0004	.28	.60	.01	.05
	Child-related education/ training	.010	1.88	.18	.19	(.13)
	Experience in current day care center	.001	1.22	.27	.11	

Table 5.9

Results of Four Regressions of Caregiver Behavior Variables
(Teacher-level, n=87)

COMFORT

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.0002	.30	.58	-.05	.03
	Child-related education/ training	.005	3.17	.08	.18	(.05)
II	Observed staff/child ratio	-.011	.24	.63	-.03	.03
	Child-related education/ training	.005	3.05	.08	.18	(.05)
III	Observed group size	-.0002	.25	.62	-.05	.05
	Child-related education/ training	.006	4.11	.05	.18	(.07)
	Experience in current day care center	-.0000	.01	.92	.01	

Table 5.10

Results of Four Regressions of Caregiver Behavior Variables
(Teacher-level, n=87)

INSTRUCT

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.002	1.59	.21	-.13	.04
	Child-related education/ training	.020	2.24	.14	.18	(.06)
II	Observed staff/child ratio	-.111	.96	.33	-.10	.03
	Child-related education/ training	.021	2.46	.12	.18	(.05)
III	Observed group size	-.002	1.55	.22	-.13	.04
	Child-related education/ training	.020	2.05	.16	.18	(.06)
	Experience in current day care center	.0002	.01	.94	.08	

Table 5.11

Results of Regressions of Caregiver Behavior Variables, Spring 1977

(Lead Teachers, n=87)

		MANAGEMENT				
	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.002	2.72	.10	.17	.03 (.06)
	Child-related education/ training	-.003	.05	.83	-.02	
II	Observed staff/child ratio	-.347	9.12	.003	-.30	.09 (.12)
	Child-related education/ training	.002	.02	.90	-.02	
III	Observed group size	.002	1.37	.25	.17	.05 (.08)
	Child-related education/ training	-.003	.05	.83	-.02	
	Experience in current day care center	-.003	1.77	.19	-.20	

Table 5.12

Results of Four Regressions of Caregiver Behavior Variables
(Teacher-level, n=87)

CORRECT

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.001	1.00	.32	.09	.03
	Child-related education/ training	-.010	1.56	.21	-.14	(.07)
II	Observed staff/child ratio	-.122	3.45	.07	-.19	.05
	Child-related education/ training	-.008	1.22	.27	-.14	(.09)
III	Observed group size	.001	.56	.46	.09	.04
	Child-related education/ training	-.010	1.48	.23	-.14	(.08)
	Experience in current day care center	-.001	.27	.60	-.12	

Table 5.13

Results of Four Regressions of Caregiver Behavior Variables
(Teacher-level, n=87)

COMMAND

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.002	3.26	.08	.19	.05
	Child-related education/ training	.007	.63	.43	.11	(.06)
II	Observed staff/child ratio	-.224	10.63	.002	-.32	.13
	Child-related education/ training	.010	1.62	.21	.11	(.14)
III	Observed group size	.001	1.56	.22	.19	.08
	Child-related education/ training	.007	.61	.44	.11	(.09)
	Experience in current day care center	-.002	3.01	.09	-.22	(.09)

the group composition measures. Teachers did less commanding in classrooms with higher staff/child ratio. There also was a tendency for more commanding in large classrooms and for teachers in higher ratio classrooms to do less correcting.

Observe (Table 5.14). The amount of time that a teacher spent observing but not actively involved with children was strongly related to both the number of children and the staff/child ratio in the classroom ($p<.01$). Lead teachers in larger classrooms tended to do more observing; conversely, teachers in higher ratio classrooms tended to do less observing.

Attention to One Child; Small, Medium, and Large Groups (Tables 5.15-5.18). Group size was a strong predictor of how caregivers distributed their attention in the classroom. The number of children with whom the lead teacher interacted was sharply related to group size. As group size increased, teachers spent significantly less time with small (2-7) and medium (8-12) groups and more time with large groups. The fact that larger group size meant less time with small and medium groups of children and more time with larger groups, suggests that large groups tended *not* to be broken into smaller activity subgroups.

Staff/child ratio also was strongly related to the number of children with whom the teacher interacted. Lead teachers in higher ratio classrooms spent significantly more time with small groups of children and less time with medium and large groups. (They also spent more time with other staff, as discussed below.)

Non-Child Activities: Center-Related Activity, Attention to Staff, and Adult-Related Activity (Tables 5.19-5.21). Only staff/child ratio, not group size, was

Table 5.14

Results of Regressions of
Caregiver Behavior Variables, Spring 1977
(Lead Teachers, n=87)

	Policy Variables	OBSERVE Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.006	10.11	.002	.32	.10 (.26)
	Child-related education/ training	-.022	1.56	.22	-.11	
II	Observed staff/child ratio	-.386	6.35	.01	-.28	.07 (.23)
	Child-related education/ training	-.014	.61	.44	-.11	
III	Observed group size	.006	9.18	.003	.32	.11 (.27)
	Child-related education/ training	-.021	1.32	.25	-.11	
	Experience in current day care center	.000	.01	.94	-.04	

Table 5.15

Results of Regressions of Caregiver Behavior Variables, Spring 1977

(Lead Teachers, n=87)

ATTENTION TO ONE CHILD

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.001	.18	.68	.04	.01
	Child-related education/ training	.026	.79	.37	.09	(.10)
II	Observed staff/child ratio	-.264	1.21	.28	-.08	.02
	Child-related education/ training	.026	.88	.35	.09	(.11)
III	Observed group size	.001	.16	.67	.04	.01
	Child-related education/ training	.026	.79	.38	.09	(.11)
	Experience in current day care center	.001	.06	.81	-.02	

Table 5.16

Results of Regressions of Caregiver Behavior Variables, Spring 1977

(Lead Teachers, n=87)

ATTENTION TO SMALL (2-7) GROUPS

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.004	8.41	.01	-.29	.10
	Child-related education/ training	.009	.40	.53	.02	(.11)
II	Observed staff/child ratio	.345	7.56	.007	.29	.08
	Child-related education/ training	.001	.01	.94	.02	(.09)
III	Observed group size	-.004	6.25	.02	-.29	.13
	Child-related education/ training	.004	.06	.81	.02	(.14)
	Experience in current day care center	.004	2.59	.11	.23	

Table 5.17

Results of Regressions of Caregiver Behavior Variables, Spring 1977

(Lead Teachers, n=87)

ATTENTION TO MEDIUM GROUPS

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.115	5.66	.02	-.26	.09
	Child-related education/ training	-.029	2.34	.13	-.13	(.14)
II	Observed staff/child ratio	-.444	7.78	.007	-.31	.10
	Child-related education/ training	-.026	1.99	.16	-.13	(.16)
III	Observed group size	-.006	9.80	.002	-.26	.18
	Child-related education/ training	-.018	.96	.33	-.13	(.23)
	Experience in current day care center	-.009	10.37	.002	-.23	

Table 5.18

Results of Regressions of Caregiver Behavior Variables, Spring 1977

(Lead Teachers, n=87)

ATTENTION TO LARGE (13+) GROUPS

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.013	32.95	.000	.54	.30
	Child-related education/ training	.031	1.96	.17	.16	(.31)
II	Observed staff/child ratio	-.608	7.90	.006	-.28	.12
	Child-related education/ training	.050	4.00	.05	.16	(.14)
III	Observed group size	.014	31.70	.000	.54	.31
	Child-related education/ training	.030	1.69	.20	.16	(.32)
	Experience in current day care center	.002	.48	.50	-.05	

Table 5.19

Results of Regressions of Caregiver Behavior Variables, Spring 1977

(Lead Teachers, n=87)

CENTER-RELATED ACTIVITY

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.003	1.28	.26	-.13	.02
	Child-related education/ training	-.001	.001	.91	-.03	(.04)
II	Observed staff/child ratio	.832	15.68	.000	.41	.16
	Child-related education/ training	-0.12	.24	.65	-.03	(.18)
III	Observed group size	-.003	1.00	.32	-.13	.02
	Child-related education/ training	-.004	.02	.89	-.03	(.04)
	Experience in current day care center	.001	.01	.91	.02	

Table 5.20

Results of Regressions of Caregiver Behavior Variables, Spring 1977

(Lead Teachers, n=87)

ATTENTION TO STAFF

	Policy Variables	Ordinary Least Squared Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.001	.21	.65	-.04	.00
	Child-related education/ training	-.002	.01	.90	-.07	(.08)
II	Observed staff/child ratio	.427	14.69	.000	.38	.15
	Child-related education/ training	-.008	.44	.51	-.07	(.21)
III	Observed group size	-.001	.17	.68	-.04	.01
	Child-related education/ training	-.003	.05	.82	-.07	(.09)
	Experience in current day care center	.001	.32	.57	.38	

Table 5.21

Results of Regressions of Caregiver Behavior Variables, Spring 1977

(Lead Teachers, n=87)

ADULT-RELATED ACTIVITY

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.002	1.74	.19	-.15	.05
	Child-related education/ training	-.018	2.43	.12	-.17	(.12)
II	Observed staff/child ratio	.022	.03	.87	-.02	.04
	Child-related education/ training	-.029	3.61	.06	-.19	(.10)
III	Observed group size	-.002	2.04	.16	-.15	.06
	Child-related education/ training	-.027	1.96	.11	-.17	(.13)
	Experience in current day care center	.000	.40	.55	-.03	

related to the amount of time teachers spent not involved with children. A higher ratio of staff to children (which usually implied more staff) meant teachers spent significantly more time in tasks which did not directly involve children--e.g., in preparation activities--and more time with other staff. This finding suggests that higher ratios give teachers more opportunity for "time-out" during the day (but do not necessarily produce less total interaction between a child and a teacher).

Discussion of Group Size and Ratio Effects

Group size and ratio were related to many of the same teacher behaviors in a pattern suggesting that larger classrooms and low ratios were disadvantageous. Larger groups sizes and lower staff-child ratios were associated with more management behavior, especially commanding, and more observing; also teachers in larger classrooms and those with lower ratios spent more time with groups of 13 or more children and less time with smaller groups. It is to be expected that the two classroom parameters were related to similar dependent measures, since group size and ratio were confounded. In this sense, it cannot be determined whether the findings represent a group size effect, a ratio effect or both.

The particularly strong relationships between the group composition measures and the TO WHOM codes may reflect teacher style or the characteristics of the instrument. That is, the way in which the FMI's were coded meant that by definition a teacher who spent more time with large groups of children had to spend less time with smaller groups. At the same time, the pattern of effects may be caused by teachers in large classrooms and in low-ratio classrooms choosing to work with larger groups. Whatever the reason, however, it appears that the actual "head count" group size is an important factor in teacher behavior.

Caregiver Qualification Measures: Child-Related Training and Experience in Current Center

Social Interaction. Whether teachers had child-related education/training was significantly related to the amount of social interaction with children. Teachers with specialized preparation tended to do more social interacting, especially the "warm" behaviors (PRAISE, COMFORT, RESPOND). Experience in current center was not significantly related to social interaction.

Management. Neither child-related education/training nor experience in current center was related to the amount of management, nor to either of the individual component codes of COMMAND or CORRECT.

Observes. Neither measure of caregiver qualifications was associated with the amount of observing a teacher did.

Non-Child Activities: Center-Related and Adult-Related Activities and Attention to Adults. There was a tendency for caregivers with child-related education/training to spend *less* time in activities that were non-child, non-center related. A teacher's experience in the current center was not related to the amount of non-child activities.

Attention to One Child, Small, Medium or Large Groups. How a teacher distributed her time was not related to her child-related education/training. A teacher's experience in the current center was related to how she distributed her attention. Longer tenure was associated with significantly less attention to medium groups, and a tendency for more attention to small groups.

Discussion of Caregiver Qualifications

Initially, four caregiver qualifications were tested in the regressions. Findings for two of these were discussed above--specialized training and experience in the current center. The other two--years of education and previous day care experience--had no effects in the regressions for lead teachers, but are discussed briefly below.

SPECIALIZATION was associated with more social interaction of all types with children and less time in activities that did not involve children. The apparently positive effects for specialized training suggests that training actually promotes beneficial behavior by the lead teacher. On the other hand, the effects might also be caused by lead teachers with specialized training having different classroom responsibilities than those without.

Although YEARS OF EDUCATION has no effects in the regressions, at the level of simple correlations its effects were similar to those for specialization. That is, caregivers with more education tended to do more social interacting with children. The problem in assessing effects for EDUCATION was its confounding with the covariables (CAREGIVER RACE and CLASSROOM SES). The confounding meant, first, that when YEARS OF EDUCATION was entered in the regressions along with the covariables, EDUCATION was never significantly related to lead teacher behavior. Second, it meant that the simple correlations for EDUCATION could not be interpreted as simply an education effect but as an effect of a complex of variables including education, race and SES.

Neither of the measures of caregiver experience was strongly related to caregiver behavior. EXPERIENCE IN THE CURRENT CENTER was associated with how the lead teacher's time was distributed. Teachers with longer experience tended to spend more time with small groups and less time with medium groups. (One caution is that experience was confounded with ratio, r=.41. Therefore, the effects might also be ratio effects, and they are consistent with the findings for ratio.) PREVIOUS DAY CARE EXPERIENCE had no effects in the regressions.*

Covariables

The covariables (race of caregiver and socioeconomic status of the children) had as strong an effect on lead teacher behavior as any of the policy variables. Both covariables were entered in each of the regression models, and they added significantly to the prediction of caregiver behavior. Since the two were highly correlated (r=.56), usually only one was significant for any dependent measure, and most often it was race of caregiver. White teachers, who also tended to be in classrooms with higher average SES, did significantly more social interacting, including more responding and praising (Table 5.22). They also did significantly less observing, and there was a tendency for less adult-related activity. In addition, they spent significantly more time with individual children, less time with medium-sized groups, and more time with other staff. In sum, the classrooms with white lead teachers and the higher-SES classrooms showed more teacher/child interaction, particularly social interaction (but with a trend toward more management as well) and less observing.

---

*Because of the relatively narrow range in the lead teachers' previous day care experience, the variable was transformed into a binary variable, with a value of "1" for some experience, regardless of amount. Comparison of teachers with some and no experience showed no significant differences. Compared to the continuous variable, the transformed binary version of PREVIOUS DAY CARE EXPERIENCE did not have any stronger relationships to caregiver behavior (or any relationships, for that matter).

Table 5.22

Simple Correlations of Spring AFI Measures and Covariables, Teachers (n=67)

	SOCIAL INTER-ACTION	MANAGE-MENT	Ob-serve	Center-Related	Adult-Related	To One Child	To Small Groups	To Medium Groups	To Large Groups	To Staff	Com-mand	Cor-rect	Ques-tion	In-struct	Praise	Com-fort	Re-spond
CLASS SES[1]																	
Correlation	-.08	.11	-.13	.13	-.15	.20	.10	-.22*	-.03	.26*	.01	.18	-.04	-.15	.16	.08	.31**
Regression Coefficient	ns	ns	ns	ns	ns	ns	ns	-.03	ns	.02	ns	ns	-.007	ns	ns	ns	ns
(and Significance)[3]								(.03)		(.01)			(.14)				
CAREGIVER RACE[2]																	
Correlation	.26*	.16	-.38**	.13	-.26*	.29**	.02	-.08	-.09	.06	.10	.18	.21*	-.05	.41**	.12	.50**
Regression Coefficient	.07	ns	-.07	ns	-.03	.07	ns	ns	ns	ns	ns	ns	.02	ns	.03	ns	.01
(and Significance)	(.01)		(.00)		(.07)	(.05)							(.01)		(.00)		(.00)

---

[1] CLASS SES was a factor score across five measures of socioeconomic status of the children in each classroom.

[2] Race was coded as "1" for black caregivers and "2" for white caregivers.

[3] Coefficients listed only if p<.15.

* p<.05.

** p<.01.

It is important to note again that since the covariables were essentially uncorrelated with the policy variables, they did not alter the relationship of the policy variables to the dependent measures. That is, although there were many significant relations between the covariables and caregiver behavior, there was virtually no interaction of the covariables and policy variables. In the tables of regression results, only the coefficients for the policy variables are given, but the R^2 accounted for by the covariables is indicated for each of the dependent measures. It is clear that the covariables frequently were responsible for most of the variance explained.

Regression Results: Aides in the 49-Center Study

Group Composition Measures

The group composition measures had moderate effects for aides. For MANAGEMENT, larger groups were significantly related to more managing, particularly with the component code COMMAND (Tables 5.23-5.25). Lower staff/child ratios were also related to more COMMAND. There were no significant relationships for group size or ratio with social interaction or its component codes (Tables 5.26-5.31). In addition, there were no relationships between GROUP/SIZE or RATIO and the measures of non-child activities--OBSERVE, CENTER ACTIVITY, ADULT ACTIVITY, or ATTENTION TO STAFF (Tables 5.32-5.35).

There were, however, significant effects for the TO WHOM codes (Tables 5.36-5.39). Larger group sizes and lower staff/child ratios were related to more time with large groups and less time with small groups.

Table 5.23

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

MANAGE

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.005	5.29	.03	.35	.13
	Child-related education/ training	.007	.12	.73	-.04	(.13)
II	Observed staff/child ratio	-.357	2.95	.09	-.28	.08
	Child-related education/ training	.002	.01	.92	-.04	(.08)
III	Observed group size	.005	5.11	.03	.35	.14
	Child-related education/ training	.006	.06	.81	-.04	(.14)
	Experience in current day care center	.001	.27	.61	.06	
IV	Observed group size	.005	4.96	.03	.35	.20
	Years of education	-.011	3.21	.08	-.29	(.20)

Table 5.24

Results of Four Regressions of Caregiver Behavior Variables

(Aides, n=42)

COMMAND

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.004	16.68	.000	.58	.31
	Child-related education/ training	.002	.04	.85	-.12	(.35)
II	Observed staff/child ratio	-.252	5.77	.02	-.40	.14
	Child-related education/ training	-.003	.09	.77	-.12	(.18)
III	Observed group size	.004	15.78	.000	.58	.32
	Child-related education/ training	-.0001	.00	.99	-.12	(.36)
	Experience in current day care center	.0001	.02	.90	-.01	
IV	Observed group size	.004	17.22	.000	.58	.34
	Years of education	-.004	1.85	.18	-.24	(.38)

Table 5.25

Results of Four Regressions of Caregiver Behavior Variables

(Aides, n=42)

CORRECT

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.001	.37	.55	.07	.01
	Child-related education/ training	.005	.15	.70	.04	(.02)
II	Observed staff/child ratio	-.105	.62	.44	-.11	.01
	Child-related education/ training	.005	.15	.70	.04	(.02)
III	Observed group size	.001	.42	.52	.07	.02
	Child-related education/ training	.006	.15	.70	.04	(.03)
	Experience in current day care center	.001	.50	.48	.11	
IV	Observed group size	.001	.16	.68	.07	.08
	Years of education	-.007	3.00	.09	-.26	(.09)

Table 5.26

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

SOCIAL INTERACTION

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.000	.00	.99	-.00	.06
	Child-related education/training	-.050	2.28	.14	-.20	(.12)
II	Observed staff/child ratio	.052	.03	.88	-.00	.05
	Child-related education/training	-.050	2.35	.13	-.20	(.12)
III	Observed group size	.001	.03	.87	.00	.12
	Child-related education/training	-.060	3.96	.06	-.20	(.19)
	Experience in current day care center	.005	1.70	.20	.06	
IV	Observed group size	.001	.12	.74	-.00	.02
	Years of education	-.005	.18	.68	-.02	(.08)

Table 5.27

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

		QUESTION				
	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.001	.25	.62	-.05	.03
	Child-related education/ training	-.014	1.03	.32	-.14	(.03)
II	Observed staff/child ratio	-.014	.01	.92	-.04	.02
	Child-related education/ training	-.011	.70	.41	-.14	(.02)
III	Observed group size	-.0002	.17	.91	-.05	.15
	Child-related education/ training	-.024	2.96	.16	-.14	(.15)
	Experience in current day care center	.008	.93	.03	.23	
IV	Observed group size	-.0004	.07	.79	-.05	.01
	Years of education	-.001	.07	.79	-.03	(.01)

Table 5.28

Results of Four Regressions of Caregiver Behavior Variables

(Aides, n=42)

RESPOND

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.0003	.51	.51	-.03	.02
	Child-related education/ training	.003	.56	.56	.11	(.20)
II	Observed staff/child ratio	.064	1.89	.18	.25	.04
	Child-related education/ training	.001	.02	.88	.11	(.22)
III	Observed group size	.0004	.45	.51	-.03	.02
	Child-related education/ training	.002	.16	.68	.11	(.20)
	Experience in current day care center	.0001	.06	.80	-.06	
IV	Observed group size	.0003	.31	.58	-.03	.01
	Years of education	.0005	.11	.75	.12	(.19)

Table 5.29

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

INSTRUCT

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.0004	.05	.82	-.01	.09
	Child-related education/ training	-.035	3.75	.06	-.26	(.13)
II	Observed staff/child ratio	-.025	.02	.89	-.07	.09
	Child-related education/ training	-.034	3.51	.07	-.26	(.13)
III	Observed group size	-.0004	.05	.83	-.01	.11
	Child-related education/ training	-.041	4.38	.04	-.26	(.15)
	Experience in current day care center	.002	.54	.47	.01	
IV	Observed group size	.0004	.04	.84	-.01	.09
	Years of education	-.004	.36	.55	-.05	(.05)

Table 5.30

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

COMFORT

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.0001	.02	.88	-.04	.00
	Child-related education/ training	.001	.03	.87	.03	(.01)
II	Observed staff/child ratio	.029	.44	.50	.13	.02
	Child-related education/ training	.0004	.11	.92	.03	(.03)
III	Observed group size	.000	.00	.98	-.04	.25
	Child-related education/ training	-.002	.23	.63	.13	(.26)
	Experience in current day care center	.002	11.09	.002	.45	
IV	Observed group size	-.001	.04	.83	-.04	.01
	Years of education	-.0004	.08	.78	-.04	(.01)

Table 5.31

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

PRAISE

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.001	.80	.38	.11	.04
	Child-related education/ training	-.005	.30	.59	-.10	(.08)
II	Observed staff/child ratio	-.001	.0001	.99	-.00	.02
	Child-related education/ training	-.007	.57	.46	-.10	(.06)
III	Observed group size	.001	1.74	.20	.11	.06
	Child-related education/ training	-.005	.38	.54	-.10	(.10)
	Experience in current day care center	.0003	.10	.75	-.06	
IV	Observed group size	.001	1.12	.30	.11	.03
	Years of education	.0001	.001	.97	.03	(.07)

Table 5.32

Results of Four Regressions of Caregiver Behavior Variables

(Aides, n=42)

		OBSERVE				
Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.005	1.60	.21	-.19	.08
	Child-related education/ training	.038	1.03	.32	.18	(.11)
II	Observed staff/child ratio	.505	1.80	.19	.24	.09
	Child-related education/ training	.040	1.15	.29	.18	(.12)
III	Observed group size	-.006	1.91	.18	-.19	.15
	Child-related education/ training	.053	1.79	.19	.18	(.18)
	Experience in current day care center	-.008	3.52	.07	-.18	
IV	Observed group size	-.006	2.55	.12	-.19	.07
	Years of education	-.007	.30	.58	-.10	(.10)

Table 5.33

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

CENTER-RELATED ACTIVITY

Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I Observed group size	.001	.03	.87	.01	.01
Child-related education/ training	.001	.001	.97	.01	(.01)
II Observed staff/child ratio	-.115	.11	.74	-.05	.01
Child-related education/ training	.002	.003	.96	.01	(.01)
III Observed group size	.000	.02	.90	.01	.03
Child-related education/ training	.000	.00	.99	.01	(.03)
Experience in current day care center	.004	.82	.37	.14	
IV Observed group size	.001	.09	.76	.01	.08
Years of education	.019	3.16	.08	.28	(.08)

Table 5.34

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

ADULT-RELATED ACTIVITY

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.001	.26	.61	-.01	.02
	Child-related education/ training	.004	.08	.78	.04	(.10)
II	Observed staff/child ratio	-.093	.49	.49	-.12	.01
	Child-related education/ training	.007	.28	.60	.04	(.09)
III	Observed group size	-.001	.43	.59	-.01	.02
	Child-related education/ training	.003	.05	.79	.04	(.10)
	Experience in current day care center	-.002	.04	.82	-.07	
IV	Observed group size	-.001	.25	.63	-.01	.01
	Years of education	.004	.81	.37	.08	(.09)

Table 5.35

Results of Four Regressions of Caregiver Behavior Variables (Aides, n=42)

ATTENTION TO STAFF

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.001	.37	.55	.06	.07
	Child-related education/ training	-.025	1.87	.18	-.20	(.19)
II	Observed staff/child ratio	-.008	.002	.97	.02	.05
	Child-related education/ training	-.028	2.35	.13	-.20	(.17)
III	Observed group size	.001	.53	.47	.06	.06
	Child-related education/ training	-.020	1.06	.31	-.20	(.17)
	Experience in current day care center	-.001	.10	.76	-.19	
IV	Observed group size	.002	.74	.40	.06	.09
	Years of education	-.010	3.06	.09	-.20	(.21)

Table 5.36

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

ATTENTION TO ONE CHILD

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.001	.02	.89	-.00	.02
	Child-related education/ training	-.027	.57	.45	-.09	(.07)
II	Observed staff/child ratio	.247	.48	.50	.10	.03
	Child-related education/ training	-.032	.84	.37	-.09	(.08)
III	Observed group size	.001	.07	.80	.00	.07
	Child-related education/ training	-.047	1.50	.23	-.09	(.12)
	Experience in current day care center	.004	.95	.34	.05	
IV	Observed group size	.001	.06	.81	-.00	.04
	Years of education	-.014	1.53	.22	-.15	(.09)

Table 5.37

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

ATTENTION TO SMALL GROUPS

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.009	8.18	.007	-.43	.21
	Child-related education/ training	.020	.50	.48	.18	(.24)
II	Observed staff/child ratio	.997	14.64	.000	.56	.31
	Child-related education/ training	.020	.62	.44	.18	(.34)
III	Observed group size	-.008	8.92	.005	-.43	.21
	Child-related education/ training	.007	.06	.81	.18	(.25)
	Experience in current day care center	-.002	.31	.58	-.06	
IV	Observed group size	-.009	9.74	.003	-.43	.20
	Years of education	-.001	.02	.90	.00	(.23)

Table 5.38

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

ATTENTION TO MEDIUM GROUPS

	Policy Variables	Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	-.002	.72	.40	-.10	.06
	Child-related education/ training	-.036	1.95	.17	-.19	(.07)
II	Observed staff/child ratio	.182	.49	.49	.08	.05
	Child-related education/ training	-.034	1.78	.19	-.19	(.06)
III	Observed group size	-.002	.65	.42	-.10	.03
	Child-related education/ training	-.023	.72	.40	-.19	(.04)
	Experience in current day care center	-.001	.11	.75	-.09	
IV	Observed group size	-.002	.39	.54	-.10	.06
	Years of education	-.012	2.01	.17	-.20	(.07)

Table 5.39

Results of Four Regressions of Caregiver Behavior Variables
(Aides, n=42)

ATTENTION TO LARGE GROUPS

Policy Variables		Ordinary Least Squares Coefficient	F	Significance of F	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
I	Observed group size	.010	5.94	.02	.35	.15
	Child-related education/ training	.056	2.08	.16	.11	(.18)
II	Observed staff/child ratio	-1.171	10.11	.003	-.44	.22
	Child-related education/ training	.055	2.25	.14	.11	(.25)
III	Observed group size	.010	5.12	.03	.35	.18
	Child-related education/ training	.072	3.12	.09	.11	(.22)
	Experience in current day care center	-.005	1.14	.29	-.09	
IV	Observed group size	.009	4.82	.03	.35	.13
	Years of education	.013	1.13	.29	.10	(.16)

Caregiver Qualifications

Three of the four qualifications variables were associated with aide behavior. YEARS OF EDUCATION, SPECIALIZATION, and EXPERIENCE IN CURRENT CENTER each had a few effects, while PREVIOUS DAY CARE EXPERIENCE had none.

SOCIAL INTERACTION and its component codes had a few relationships to aides' qualifications. There was a trend for aides with specialized training to do less social interacting and less instructing. YEARS OF EDUCATION was not associated with amount of social interacting. Aides with more experience in the center, however, did significantly more praising and comforting.

MANAGEMENT and its component codes was not significantly related to any qualifications. There was a trend for aides with more education to do less correcting and less managing.

The effects for the non-child activities were small. There was a trend for less observing by aides with more center experience. The amount of time in center activity or staff activities was weakly related to aides' qualifications. There were trends for less center activity among aides with specialization and more activity among aides with more education. The amount of non-child, non-center activity was not related to aide qualifications. There were no significant relationships between aide qualifications and the TO WHOM codes.

Covariables

The effects of the covariables for aides were small, much smaller than the effects for teachers (Table 5.40). The classroom SES was never a significant predictor

Table 5.40

Simple Correlations of Spring API Measures and Covariables, Aides (n=42)

	Social Interaction	Management	Observe	Center-Related	Adult-Related	To One Child	To Small Groups	To Medium Groups	To Large Groups	To Staff	Command	Correct	Question	Instruct	Praise	Comfort	Respond
CLASS SES[1]																	
Correlation	.12	-.06	-.05	.03	-.16	.09	.13	.10	-.15	.27	-.19	.07	.02	.04	.13	.10	.30**
Regression Coefficient and Significance[3]	ns[3]	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
CAREGIVER RACE[2]																	
Correlation	.25	-.04	-.17	.07	-.29	.22	-.03	.09	-.17	.32*	-.15	.06	.05	.20	.20	.04	.41**
Regression Coefficient	.06	ns	-.06	ns	-.02	ns	ns	ns	ns	.04	ns	ns	ns	.03	ns	ns	.01
(and Significance)	(.11)		(.14)		(.12)					(.08)				(.12)			(.05)

---

[1] CLASS SES was a factor score across five measures of socioeconomic status of the children in each classroom.

[2] Race was coded as "1" for black caregivers and "2" for white caregivers.

[3] Coefficients indicated only if $p<.25$.

* $p<.05$.

** $p<.01$.

for aide behavior. Race of caregiver had a significant effect only for RESPOND; and, as was true for teachers, white aides more often responded to children. The pattern of other nearly-significant effects was consistent with that for teachers, with more social interacting and less observing, and more time in staff activities for white aides.

Discussion of Findings for the 49-Center Study

The findings of the regression analyses for 49-center lead teacher and aides are summarized in Table 5.41. For the *group composition* measures, there was sufficient consistency in the findings across the two samples of caregivers to discuss them as a whole; that is, there was little evidence of an interaction between caregiver role in the classroom and the group composition measures. For teachers and aides, the size of the classroom was strongly related to the size of activity subgroups: larger classrooms meant caregivers dealt more often with large groups (13 or more children). Also, the size of the classroom was related to what caregivers did, with more observing, more management interaction, and less social interaction. Thus, in smaller classrooms, it seems likely that a child received more direct interaction with the caregiver.

Regarding the size of activity subgroups, the effects for ratio followed a similar pattern to the effects for group size. Higher ratios were associated with more time with small groups. On the other hand, ratio affected *what* caregivers did in a different pattern than group size did. Higher ratios were associated with more non-child-interaction activities as well as less observing and managing. Caregivers in high-ratio classrooms tended to spend more time in center-related activities such as preparation and in interaction with other staff.

Table 5.41

Summary of Significant Regression Results[1] for Spring AFI

LEAD TEACHERS

	SOCIAL INTERACTION	MANAGE	Observe	Center Activity	Adult Activity	Attn. to Staff	Attn. to Child	Attn. to Small Group	Attn. to Med. Group	Attn. to Lg. Group	Command	Correct	Question	Praise	Comfort	Respond	Instruct
Group Size		(+)	+					-	-	+	(+)					-	
Staff/Child Ratio		-	-	+		+		+	+	-	-	(-)					
Specialization	+				(-)				(-)				(+)	+	(+)	(+)	(+)
Experience in Current Center								(+)	-		(-)						
Years of Education																	
Previous Day Care Experience																	
CG Race (+ = White) Class SES	+		-		(-)	+	+		-				(-)	+		+	

AIDES

	SOCIAL INTERACTION	MANAGE	Observe	Center Activity	Adult Activity	Attn. to Staff	Attn. to Child	Attn. to Small Group	Attn. to Med. Group	Attn. to Lg. Group	Command	Correct	Question	Praise	Comfort	Respond	Instruct
Group Size		+						-		+	+						
Staff/Child Ratio		(-)						+		-	-						
Specialization	(-)					(-)				(+)							(-)
Experience in Current Center			(-)										+		+		
Years of Education		(-)		(+)		(-)						(-)					
Previous Day Care Experience																	
CG Race (+ = White) Class SES	(+)		(-)		(-)	(+)	+									+	+

---

[1] Results noted were significant at p<.05; results significant at .05<p<.15 shown in parentheses.

In both samples, there were a few significant findings that suggested positive effects for more caregiver qualifications. However, there was little consistency across the teacher and aide samples. The caregiver variables, SPECIALIZATION and EXPERIENCE in current center, had scattered effects in both samples. For SPECIALIZATION, teachers with specialized training tended to do more social interacting. SPECIALIZATION had few effects for aides, possibly because so few aides had specialized training. Those aides with specialized training tended to spend more time with large groups of children and less time in staff activities, which suggests more direct responsibility for children. However, they also tended to do less instructing. EXPERIENCE in current center was associated with more attention to small groups and less commanding for teachers, and with more "warm" interactions with children (QUESTIONS, COMFORTS) and less observing for aides. YEARS OF EDUCATION had some effects, but for aides only. Aides with more education tended to manage less frequently and to spend more time in non-child activities. Since caregivers in supervisory roles (e.g. center directors) also tended to exhibit these behaviors, this may be a reflection of aides with more education being put in more of a supervisory role for the classroom. Previous day care experience had no effects in either sample.

In the following sections, three additional sets of analyses are discussed: tests of other independent measures, analyses of fall AFI data, and analyses of APS data. These analyses suggest general consistency in the findings for all of the independent measures, with particularly good stability for the group composition variables.

Regression Results: Fall AFI Analyses in the 49-Center Study

Lead Teachers

In the parallel regression analyses of the fall 49-center AFI data, virtually none of the effects contradicted those in the spring data (Table 5.42). Where the effects of a predictor were significant for both fall and spring, the results were generally similar in direction. Typically, however, a significant effect in one data set was matched by a nonsignificant effect in the other data set (although the simple correlations were virtually always in the same direction). Group size, ratio and the covariables had relatively stable effects across the fall and spring data sets. The caregiver qualifications variables had scattered effects at both timepoints, and few overlapping effects. Also the regressions for the TO WHOM codes generally were more similar at fall and spring than those for the WHAT codes, because of the strong and consistent effects for the group composition measures.

Aides

The regressions carried out on the aide samples were generally less consistent across timepoints, compared with those for teachers (Table 5.43). The absence of strong effeccts in either the fall or spring aide samples helped to create the picture of inconsistent results at fall and spring. The relatively small sample sizes probably contributed to the lack of consistent results. Besides the covariables, only ratio showed any consistent effects across the timepoints, and the covariables were much stronger predictors for aides in the fall. As was true for lead teachers, the regression results for the TO WHOM codes were more parallel at fall and spring than the results for the WHAT codes.

Table 5.42

Comparison of Fall and Spring Significant Regression Results for Teacher Behavior[1]

(49-Center Lead Teachers: n=83 at Fall; n=87 at Spring)

	SOCIAL INTER-ACTION		MANAGE-MENT		Center Activity		Ob-serve		Adult Activity		To One Child		To Small Group		To Medium Group		To Large Group		To Staff		Com-mand		Cor-rect		Praise		Com-fort		Re-spond		In-struct		Ques-tion	
	F*	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S
Group Size[2]			+		-			+					-	-	-	-	+	+				+	+							-	(+)			
Staff/Child Ratio			-	-		+		-					+	+		-	-	-	+	+		-	-	-					+					
Child-related Education/ Training		+	+							(-)						(+)					(+)					+		+		+		(+)	(+)	(+)

---

[1] Signs represent significant regression coefficients, p<.15; signs in parentheses represent coefficients significant at .10<p<.15.

[2] Results are presented only for the three independent measures that had significant effects.

*F=Fall; S=Spring.

Table 5.43

Comparison of Fall and Spring Significant Regression Results for Aide Behavior[1]

(49-Center Aides: n=39 at Fall; n=42 at Spring)

	SOCIAL INTER-ACTION		MANAGE-MENT		Center Activity		Ob-serve		Adult Activity		To One Child		To Small Group		To Medium Group		To Large Group		To Staff		Com-mand		Cor-rect		Praise		Com-fort		Re-spond		In-struct		Ques-tion	
	F*	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S
Group Size[2]				+										−	−		+	+				+												
Staff/Child Ratio			−	−			+						+	+			−	−			−	−									+			
Child-related Education/ Training		(−)			(−)		(+)						(+)				−										−					−		

---

[1] Signs represent significant regression coefficients, p<.15; signs in parentheses represent coefficients significant at .10<p<.15.

[2] Results are presented only for the three independent measures that had significant effects.

*F=Fall; S=Spring.

Neither group size nor the caregiver variables, however, had particularly stable effects.

Discussion

It should be noted that there are multiple possible explanations for inconsistencies between the fall and spring results. For example, different data collection procedures might be responsible for the differences. In the spring, each caregiver was observed on two mornings, by two observers, one black and one white. In the fall, each caregiver was observed on one or two mornings, by one observer, usually of the same race. Thus, the fall results are more likely to be confounded with observer differences, particularly observer race. Fall-spring differences also might result from fallible measures, both independent and dependent. A third possible reason for differences is that changes may have occurred in the day care classrooms over the year. It is possible that variables such as group size operate differently in classrooms in October, when classrooms were still getting organized, and in classrooms in April. If so, fall-spring differences might reflect actual differences in relationships between policy variables and specific caregiver behaviors at different times in the day care year, although the same policy variables seemed to have generally positive effects at both time points. There were some notable differences in the fall in the intercorrelations of the independent measures, particularly for aides. Among aides, there was more confounding among the independent measures at fall than at spring, particularly between the experience variables and the classroom parameters.

Regression Results: Caregivers in the Atlanta Public School (APS) Centers*

A major substudy of the NDCS involved eight centers operated by the Atlanta Public Schools. The APS centers provided a strong test of the effects of the policy variables, first because of the homogeneity of the child population and caregivers in the centers (all were black and similar in socioeconomic status) and second because the APS experiment involved random assignment of children to classes within centers. It was hypothesized that the effects of the policy variables would be similar in direction to the 49-center effects but stronger in the APS sample.

---

*For more detailed discussion of the APS findings and the APS experiment, see Goodrich, N. N. (1980).

References

Abt Associates Inc. "An Overview of the Study." *National Day Care Study First Annual Report 1974-1975*, Volume I. Cambridge, Mass.: Abt Associates Inc., 1976.

Abt Associates Inc. "National Day Care Study Phase II Research Report." Unpublished. April 1977.

Bache, W.L. "Comparing Alternative Measures of Classroom Composition." *National Day Care Study Measurements and Methods*. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980a.

Goodrich, N.N. "The Effects of Day Care in Eight Atlanta Public Schools Day Care Centers." *National Day Care Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980b.

Medley, D.M. and Mitzel, H.E., "Measuring Classroom Behavior by Systematic Observation." In N.L. Gage, *Handbook of Research on Teaching*. Chicago, Ill.: Rand McNally and Co., 1963.

Singer, J.D., Affholter, D.P., and Goodrich, R.L. "Variance Components and the Dependability of Measures Used in the National Day Care Study." Unpublished. 1978.

Stallings, J. and Broussard, D. *Final Report National Day Care Cost-Effects Study: Spring 1977 (Phase III) Data Collection*. Menlo Park, Calif.: SRI, 1977.

The Child-Focus Observation Effects Analysis

David B. Connell

TABLE OF CONTENTS

Page

CHAPTER ONE: INTRODUCTION TO THE NDCS 113

NDCS Phase III Design 115
  49-Center Study 116
  APS Study 118

CHAPTER TWO: THE CHILD-FOCUS INSTRUMENT 121

Phase III Sample and Procedures 129

CHAPTER THREE: OBSERVER TRAINING AND RELIABILITY 132

Training of Observers 133

Reliability of Child Observation Codes 136

CHAPTER FOUR: SELECTION AND CONSTRUCTION OF DEPENDENT MEASURES 142

Reflection/Innovation 144

Verbal Initiative 144

Cooperation/Compliance 145

Noninvolvement (No Task) 145

Aimless Wandering 146

Task Persistence 146

Patterns of Interaction 147

Classroom Activity Balance 150

Activity Transition 151

CHAPTER FIVE: SELECTION OF INDEPENDENT MEASURES 152

CHAPTER SIX: RESULTS OF CFI ANALYSES 161

Experimental Study of Ratio Effects 162

Regression Analysis 163

Discussion of Regression Model 168

Determinants of Rare But Important Events 189

Child Behavior in Structured Situations 193

CHAPTER SEVEN: CHILD OUTCOMES AND CHILD CARE POLICY 201

References 207

CHAPTER ONE: INTRODUCTION TO THE NDCS

The National Day Care Study (NDCS) was a four-year study of center-based preschool day care. It was initiated in 1974 by the Office of Child Development (now Administration for children, Youth and Families). The major objective of the NDCS was to determine the impact of variations in staff/child ratio, number of caregivers, group size and staff qualification on both the development of preschool children and the costs of center care.

The National Day Care Study (NDCS) was undertaken to provide empirical information in three areas of major policy concern: (1) the quality of day care provided, (2) the per-child cost of the day care provided; and (3) the potential quality-cost trade-offs associated with alternative formulations of federal day care regulations. The study focused on the largest group of children receiving federally subsidized care--preschool children (aged 3-5)--and on the day care settings in which most of these children are found--urban day care centers serving low-income families. The study also focused on program characteristics that have long been considered key determinants of quality and cost in center care--staff/child ratio, group size and caregiver qualifications. The mandate of the NDCS was to examine the impact of variations in these and other regulatable characteristics on the quality and cost of care received by preschool children.

The NDCS addressed the following major policy questions:

- How is the daily experience and consequent development of preschool children in day care centers affected by variations in staff/child ratio, group size, caregiver qualifications and other regulatable center characteristics?

- How is the per-child cost of center-based day care affected by variations in staff/child ratio, group size, caregiver qualifications and other regulatable center characteristics?
- How does the cost-effectiveness of center-based day care change when adjustments are made in staff/child ratio, group size, caregiver qualifications and other regulatable center characteristics?

This document reports analytic findings relevant to the first policy question, specifically, how is preschool child activity and behavior in day care centers related to variations in key program characteristics (group size, ratio, and caregiver qualifications)? The activities and behavior of children in the day care setting is possibly the most crucial category of measurement in studying quality of day care since it directly reflects the experiences that are presumed to influence the social, emotional and cognitive growth of children. Relationships between the key NDCS program characteristics and child activity in the centers, therefore, might suggest regulatory approaches that positively influence the quality of care for children.

Selection of observation techniques as opposed to other data collection methods, such as rating scales (by teachers or observers) and standardized tests, is related to the goals and independent variables in the NDCS. The NDCS goals were to relate various classroom parameters to the experience of children in the day care environment and to assess the long term impact of those experiences. At the outset of the NDCS, a number of standard tests and ratings were proposed to measure various social skills of children in day care. For various reasons, each of those tests were rejected, primarily on psychmetric grounds (Stallings and Broussard, 1977). In addition, it was apparent that the links between the associated social skills and the day care classroom vriables were indirect at best, and in many cases were justified only by conjecture. Direct observation of children offered the best possibility of linking meaningful child behaviors in the day care center directly to descriptors of the center itself. Such procedures have strong intuitive appeal. The connection between data and phenomena is usually direct. Observations provide researchers with some of the indicators of quality that are available to parents when choosing a day care center for their child, i.e., impressions of the degree to which the center provides stimulating social interaction among children and between adults and children, and the extent to which it elicits cooperative, creative and verbal intellectual activity on the part of the child. Natural observations avoid the artificiality that opens many laboratory studies to the charge that their findings have nothing to do with real-world behavior. Use of such observations in the NDCS exemplifies the "ecological" approach to the study of child development urged by some of the field's most prominent spokesmen, notably Urie Bronfenbrenner.

NDCS Phase III Design

The NDCS was conducted in 57 day care centers in three sites.* Atlanta, Detroit and Seattle were chosen as the study sites, and a total of 57 centers subsequently were selected for participation. The centers were selected for high or low values of staff/child ratio, group size and staff education. The first year following site selection was a study of naturally existing relationships between regulatable center characteristics and measures of classroom process, including caregiver observations. This phase of the study was used to (1) formulate initial hypotheses about relationships among regulatable center characteristics, classroom process and developmental outcomes; and (2) refine the measures of regulatable characteristics, classroom process and developmental outcomes to be used in in the third and final year of the study (Travers, Coelen, and Ruopp, 1977).

---

*Site and center selection were carried out during Phase I of the study, July 1974 to September 1975 (Abt Associates, 1976).

Phase III (October 1976 to September 1977) was designed to answer the study's three major policy questions (Travers, Coelen, and Ruopp, 1977). The Phase III investigation had two components: a quasi-experiment conducted in 49 of the centers in all three sites, and a randomized experiment conducted in eight centers operated by the Atlanta Public Schools (APS). The eight APS centers were not included in the 49-center sample. In both studies, selected center characteristics were altered systematically, permitting measurement of the effects on caregiver behavior associated with such changes.

49-Center Study

The quasi-experiment in 49 of the centers was a comparison of three groups of centers (Figure 1.1). Group I (the "treatment" group) consisted of 14 centers which had low observed staff/child ratios (1:9.1) in Phase II, and whose ratios were increased to 1:5.9 in Phase III.* Effects of this treatment on caregivers and children were compared with results from a matched group of 14 untreated centers with naturally low-ratio (1:9.1), labeled Group II, and a group of 21 untreated centers with naturally high ratio (1:5.9), labeled Group III. The question of central interest in the quasi-experiment was whether the experimentally-induced increase in staff/child ratio would produce different

---

*Note that, in conformance with HEW directives, manipulations consisted only of making low ratios higher. The Group I treatment simulates one potential effect of full enforcement of FIDCR under Title XX--namely an increase in ratios in centers serving publicly funded children but currently operating below FIDCR ratios.

Figure 1.1

DESIGN OF THE 49-CENTER QUASI-EXPERIMNENT

Group I - Treated centers (Observed mean ratio for 14 centers = 1:9.1 in Phase II; ratio raised to 1:5.9 in Phase III)

Group II - Untreated low-ratio centers (Observed mean ratio for 14 centers = 1:9.1)

Group III - Untreated high-ratio centers (Observed mean ratio for 21 centers = 1:5.9) caregiver behaviors than were seen in the matched low-ratio centers (Group I vs. Group II). A supplementary question was whether the treated and untreated high-ratio centers (Group I and II) looked different. That is, would the experimental increase in ratio eliminate most or all differences between centers that previously operated at different ratios, or would differences in outcomes continue to exist, presumably because of other center characteristics that normally accompanied high ratios but were unaffected by the experimental increase in ratio?

Ratio was chosen for manipulation because of its critical policy relevance; manipulation would reduce any confounding between ratio and other center characteristics, permitting clearcut assessment of its effects. Group size, caregiver experience and years of education were distributed as evenly as possible across the three experimental groups, so that the effect of ratio could be clearly separated out. No attempt was made in the quasi-experiments to alter natural variations in age-grouping. The three sets of ratios applied to classrooms that served primarily three- and four-year old children. In some centers, three-year olds were clearly separate from four-year olds; in others, the two ages were mixed in the same classroom.

APS Study

The APS Study was an eight-center, 29-classroom experiment in which children were randomly assigned, within centers, to classrooms that differed systematically in level of staff education and staff/child ratio (Figure 2). Group size and caregiver experience were distributed as evenly as possible across the three experimental groups. Twelve of the experimental classrooms served three-year old children, and seventeen served four-year olds. This design made possible a relatively clearcut assessment of the effects and interactions of staff education and staff/child ratio for children of different ages (three- and four-year olds).

Figure 1.2

DESIGN OF THE ATLANTA PUBLIC SCHOOLS (APS) EIGHT-CENTER EXPERIMENT

	High Ratio (Observed Mean Ratio = 1:5.4)	Low Ratio (Observed Mean Ratio = 1:7.4)
High Staff Education	4 classrooms	4 classrooms
Medium Staff Education	7 classrooms	4 classrooms
Low Staff Education	6 classrooms	4 classrooms

Staff in the APS centers fell into three distinct categories of educational background. First, center directors (who were required to work in classrooms as well as to function as directors) had bachelor's degrees; most also had master's degrees. Second, lead teachers were graduates of the Atlanta Area Technical School (AAT) two-year, post-secondary training program in day care or had completed at least two years of college. Third, aides generally had high school diplomas (or an equivalent such as the G.E.D.); the majority of aides had also completed the 60-hour, state-required training courses in day care offered through AAT. As shown in Figure 1.2, persons at these three levels of education were assigned to be lead teachers in the experimental APS classrooms--some in classes with relatively high staff/child ratios, others in classes with lower ratios. Thus, ratio and education were crossed in a two-way factorial design. Children were then randomly assigned within centers to these experimentally organized classes. Random assignment, together with the fact that the children served by APS centers were unusually homogeneous in ethnic and socioeconomic background (virtually all were black children from low-income families) minimized any confounding of center characteristics and children's background characteristics.

The two Phase III components addressed the same questions but had designs with different experimental strengths and weaknesses. Because the 49-center study included a large and diverse group of centers in three different sites, its results, if uniform across the sample, were likely to be widely generalizable; however, the diversity of the sample also posed challenges for analysis and interpretation. The APS study provided a greater degree of experimental control and afforded more safeguards against confounding of center characteristics with characteristics of the children, families or communities served. However, the generalizability of its results was potentially limited by the homogeneity of the sample.

CHAPTER TWO: THE CHILD FOCUS INSTRUMENT

Behavior of young children in day care is varied and volatile--much more so, for example, than behavior of children in elementary school settings. The NDCS required an observation instrument and analytic approach that could do justice to this complexity, yet yield a manageable set of behavior descriptors that reliably characterized children, classes or centers along dimensions relevant for assessing quality. The Child-Focus Instrument (CFI), used in the NDCS for naturalistic observation of children, was based on the Child Observation System developed by Elizabeth Prescott (Prescott, Jones, Critchersky, Milich, and Hazelhoef, 1975). SRI selected the Prescott instrument after reviewing several alternative systems and conducting field tests of the most promising candidate instruments during Phase I (Stallings, Wilcox and Travers, 1976) The Prescott instrument was attractive because it had been developed specifically for preschool children in day care settings and because it had been used for research purposes quite close to those of the NDCS (Prescott, et al., 1975). The system includes a large number of behavior codes, many of which are highly specific and have a fairly high degree of face validity and objectivity. SRI was able to train observers to high levels of accuracy for almost all codes, both in initial field testing and in subsequent use during Phases II and III.

The CFI was modified several times in the course of the NDCS; the version described here is the one used in Phase III (Stallings and Broussard, 1977). Each child observation consists of a twenty-minute period, broken into 100 twelve-second coding intervals. Observers were provided with timers that click every twelve seconds and instructed to record the behavior of a preselected focus child at the time of each click. Each record or frame has three parts:

- a section containing one of 50 codes characterizing the child's principal behavior during the 12-second coding interval. These include 37 activity codes, used when the child engages in some form of overt action, and 13 "receives" codes, used when the major event during the coding interval is an initiative directed toward the child by some other person, e.g., a request, praise or correction. Additional codes accompany some of the "receives" codes to indicate whether the child's response is appropriate.

- a section containing four object codes (adult, child, group of children, or environment), indicating the person(s) or thing(s) toward which the focus child's attention is directed; and

- a section containing three activity continuity codes, indicating whether the child's behavior is a new activity, an old activity, or no identifiable activity at all.

Table 2.1 lists the codes and shows their relative frequencies of occurrence in the Phase III data, i.e., their frequencies as percentages of all 725,000 frames recorded in fall and spring.* Definitions of the more important codes are provided immediately below. Description of the data base and data-gathering procedures appear in the following section.

Many of the CFI codes shown in Table 2.1 are specific and self-explanatory. However, some of the most frequently occurring codes (e.g., "shows closed, structured activity") are broader and require some explication. The following definitions of the most common activity and "receives" codes have been excerpted from SRI's training manual:

---

*Frequencies of the activity continuity codes indicating old vs. new activities are not shown directly in the table. By a procedure outlined in the later section on construction of dependent variables, these two codes were used to compute the duration of the child's longest single activity during the 20-minute observation period. The latter figure is shown in the table.

Table 2.1

FREQUENCIES OF CHILD OBSERVATION CODES[a]

(FALL, 1976 AND SPRING, 1977)

A. Activity Codes	Percent of All Frames
Group closed, structured activity	21.1
Group open, expressive activity	13.2
Monitors environment (looks, watches)	11.9
Gives opinions	8.0
Wanders aimlessly, does nothing	5.3
Group passive activity	4.8
Moves with purpose	3.1
Individual open, expressive activity	2.9
Adds prop or idea	2.8
Considers, contemplates problem	1.7
Individual closed, structured activity	1.5
Gives orders, directs others	1.0
Intrudes playfully	0.9
Asks for attention	0.9
Selects activity (with others)	0.6
Shares, helps	0.6
Asks for information	0.4
Asks for turn	0.3
Selects activity (alone)	0.3
Isolates self	0.3
Asserts rights	0.3
Cries	0.2
Sees pattern, solves problem	0.2
Intrudes hostilely, bullies	0.1
Hostilely asserts rights, anger	0.1
Hostile exchange	0.1
Avoids, withdraws	0.1
Individual passive activity	0.1
Asks for assistance, help	0.1
Offers sympathy, comfort	0.1
Asks for comfort	0.1
Intrudes unintentionally	0.1
Experiences rejection	0.1
Quits activity after frustration	<0.1
Angry reaction to frustration	<0.1
Experiences accident	<0.1
Temper tantrum	<0.1

Table 2.1 (continued)

B. "Receives" Codes[b]	Percent of All Frames	Percent of Appropriate Response
Receives general comments	5.1	86.5
Receives information, guidance	4.7	87.1
Receives demands, requests	4.0	82.3
Receives request to play, share	0.5	63.1
Receives rules, corrections	0.4	70.4
Receives punishment, threats	0.3	47.7
Receives praise	0.3	82.6
Receives playful intrusion	0.9	
Receives comfort	0.2	
Receives hostile intrusion	0.1	
Receives unintentional intrusion	0.1	
Receives physical punishment	<0.1	
Receives rejection	<0.1	
C. Object Codes		
Attention to adult	27.3	
Attention to child	23.0	
Attention to group	7.8	
Attention to environment	41.9	
D. Activity Continuity Codes		
Longest Activity[c]	54.8	(11 Minutes)
Not involved in activity	7.3	

---

[a] Code frequencies are shown as a percentage of all observation codes (excluding structured situation observations). For both behavior and object codes, the total number recorded was approximately 725,000.

[b] The "receives" codes indicate initiatives by others toward the child. The column headed "Percent of All Frames" shows the frequencies of these codes as percentages of all 725,000 codes. The column headed "Percent of Appropriate Responses" indicates how often children responded appropriately to selected initiatives.

[c] The "longest activity" code is computed as a percentage ratio of the duration of the longest activity to the total duration of the observation period. Since the observation period usually lasted 20 minutes, the longest activity of the typical child lasted .548 x 20 minutes, or 11 minutes.

Participates in group activity--closed, structured: Focus child and others are involved in an activity that has a goal, clear guidelines for carrying out the task, and a defined beginning and end. Focus child's participation in adult-directed group activities is coded here. (The presence of other children in the activity differentiates this code from individual structured activity, discussed below.) Examples: child is part of a group playing musical chairs; or child and a friend are working together to clean off the table.

Participates in group activity--open-ended, expressive: Focus child participates with others in a mutual experience that has no goal, no external guidelines or defined point of completion; the structure of the activity is determined by those involved, not by the materials. (The presence of other children in the activity differentiates this code from individual open-ended activity, discussed below.) Examples: Child is playing with other children in the block corner; or child and another child are swinging alongside each other, making a game of who can swing higher.

Monitors environment (looks, watches): Focus child's attention is obviously directed at other people or things. This code is not used for listening. The focus child may be either in or out of an activity. The object code used with this code identifies the focus of the child's attention. Examples: Child stands apart from a group of children, watching them play; or child is playing at the block table; his attention is directed to an adult cleaning up some spilled paint.

Gives opinions, preferences, information, comments: Focus child initiates statements about his own likes, dislikes or preferences. This code also includes information and comments initiated by the focus child (not in response to a question). Examples: "I went on a picnic yesterday;" or "Johnny is my best friend."

Does nothing, wanders: Focus child wanders around center with no apparent purpose to his movement. He may be sitting or standing doing nothing, looking around the area with no apparent focus.

Examples: Child wanders from sandbox to slide and then to doll corner, not concentrating on anything or anyone.

Participates in group activity--passive attention: Focus child is part of a group that is involved in an activity which requires no visible response, but does require concentration or thought. (The presence of other children in the activity differentiates this code from monitoring the environment.) Examples: Child and other children are watching a puppet show; or child is part of a group that is watching TV; or child is part of a group to which an adult is reading a story.

Moves with purpose: This code is used when the focus child is going from one activity to another or whenever it seems evident that there is some goal to his movement. Examples: Child has just finished gluing on a piece of paper; he heads for the bathroom to wash his sticky hands; or, child notices that a swing is free and runs across the yard toward it.

Individual open-ended, expressive activity: Focus child is involved in an activity that has no defined goal, external guidelines, or defined point of completion; the structure of the activity is determined by the child. Other children do not share in this activity with the focus child--he is alone. Examples: Child is playing with blocks; or child is dancing alone to a record.

Adds a different prop or new idea: Focus child adds variety to his activity. He uses a different toy or prop from the one he was using previously in the same activity, or he uses the same prop in a different way. This code is also used when the focus child resumes play with an article that he used formerly in the same activity. Examples: child adds a different color to his painting; or child is washing dishes in the doll corner, then picks up a doll and washes it.

Considers, contemplates, tinkers: Focus child considers before making a selection of materials. Focus child tries out an object, looks at it, moves it, examines it, manipulates it. Focus child struggles with a problem, attempting to solve it. Examples: Child carefully examines a truck, checking out each moving part; or child pulls on cargo net and watches how the net moves in response to his pull.

Individual structured, closed activity: Focus child is involved in an activity which has a goal, clear guidelines. Other children do not share in this activity with the focus child. Examples: Child is stringing beads for a necklace; or child is working on a puzzle; or child is alone at a table, grating cheese for a pizza.

Receives orders or minor behavioral corrections: Focus child receives commands with which compliance is expected. This code also includes orders to maintain smooth operation of the center and minor behavioral corrections. Examples: Adult tells child to put books away; or another child says to focus child, "Let me have the trike now."

Receives information/help with a task: Focus child receives instruction, materials, or assistance related to his task or the solution to his problem. This code includes verbal and nonverbal assistance or demonstration. Also included in this code are preliminary directions and review of an activity. Examples: Child is having difficulty completing his puzzle and the teacher shows him where the piece goes; or adult is telling focus child how to clean paint brushes.

Receives general comments, questions: Focus child is asked for information or receives comments of a general nature. Examples: Adult says to child, "Today is Johnny's birthday"; or another child tells focus child, "My grandma made this dress."

As shown in Table 2.1, recorded frequencies of the behavior codes varied widely. In Phase III, the eleven activity codes and three "receives" codes defined above occurred more than once per 20-minute observation (i.e., more than one percent of the time). Most analyses reported in later sections are based on these common codes and combinations thereof. However, many codes of psychological interest occurred rarely--a few times per thousand frames, or less. Many of the latter were events that are potentially important as indicators of harm; a few were potential indicators of benefits of day care. Examples include the codes "cries," "isolates self," "refuses to comply," "experiences accident," "shares or helps," and a number of codes indicating anger or hostility.

There are several possible reasons for the low frequencies of these events. One is that frequencies of events recorded with a time-sample instrument such as the CFI depend partially on the durations of those events. If psychologically important events are brief, they will appear in few frames or be missed altogether. For example, a hug or a slap may last less than a second. When such events occur they are likely to be very important to the children involved and memorable for an adult who happens to witness them. Yet a time sampling record of a day in which one such event occurs will show that the event occupied a tiny fraction of the observation period. In contrast, more commonplace activities such as game playing or story telling would occupy a much greater portion of the day.

A second reason concerns limited opportunities for children to display behaviors that meet the definitions of relevant codes. For example, sharing, taking turns and helping with minor tasks are routinized in most centers. Routinized prosocial behavior is coded as a form of group activity, or as compliance with adult requests, rather than as voluntary helping or sharing, accounting for the rarity of this particular code. Similarly, most centers are organized to prevent conflict and to terminate it quickly when it occurs. To the degree that they succeed, "opportunities" for conflict are limited, and associated codes are rare.

The rarity of important codes was addressed in two ways. First, rare codes from the natural observations were analyzed separately from more frequent codes, using a form of statistical analysis more appropriate for rare events than ordinary regression. Second, in addition to natural classroom observations, children were observed in structured situations designed to provide greater opportunity for voluntary prosocial behavior such as helping and sharing. Results of both approaches are presented in this report following discussion of the main analyses and results.

Phase III Sample and Procedures

The study design called for each child to be observed four times for a total of eighty minutes in both fall and spring--three times in natural situations (primarily free play and teacher-directed activity) and once in a pair of structured situations. In the spring, natural observations were conducted by two different observers for each child--generally one black and one white observer --in order to permit analytic separation of actual behavioral differences among children from differences in perspective among observers. SRI was able to implement the design with substantial success, as the following data indicate.

Approximately 8,300 twenty-minute observations of target children were completed by SRI's observers. The distribution of observations between time points and between natural (classroom) and structured observations is shown in Table 2.2. Numbers of children and classrooms observed are also shown in the table. Of 1,108 children observed in the spring, 1,086 had been observed in the fall. At both times, the sample was approximately evenly divided among Atlanta Public School centers, Atlanta centers outside the public schools, Detroit centers and Seattle centers.

Table 2.2

PHASE III CHILD OBSERVATION SAMPLE

	Fall 1976	Spring 1977
Natural (Classroom) Observations		
Number of Observations	3,987	3,177
Number of Children	1,310	1,108
Number of Classrooms	117	116
Structured Observations		
Number of Observations	642	523
Number of Children	1,284	1,046

In both fall and spring, natural observations took place in four general types of situations: free play, adult-directed activity (including both individual and group activities, with the latter predominant), routine center activities (cleanup, snack, toileting, etc.) and multiple activities--combinations of two or more of the preceding types occuring within one twenty-minute observation. By design, free play and teacher-directed activities were observed most frequently. As shown in Table 2.3, about 38 percent of fall observations and 41 percent of spring observations took place during free play periods; 42 percent of fall observations and 41 percent of spring observations occurred during teacher-directed activities. Since the dynamics of the group can change dramatically across these general types of situations, separate analyses were conducted for data from free play and teacher-directed periods.

Table 2.3

CHILD FOCUS OBSERVATION SITUATIONS
(percentages)

	Fall 1976	Spring 1977
Free Choice	38%	41%
Center routine	10	9
Adult-directed individual	4	3
Adult-directed group	38	38
Principle types	10	9

CHAPTER THREE: OBSERVER TRAINING AND RELIABILITY

Of all threats to the validity and reliability of observation instruments, the one that has received the most attention in the psychological literature is distortion of results due to differences in observer perspective. Characteristically, considerable effort is devoted to training observers to high criteria of agreement, and often, when such standards are achieved, the researcher assumes that his or her measures are trustworthy. Although, as shown in the next section, the importance of observer effects is usually overated and high observer agreement is no guarantee that measures are dependable, observer effects nevertheless deserve careful attention.

The first line of defense against observer effects lies in training. SRI recruited and trained observers carefully and tested their performance on selected videotaped samples of behavior before and after sending them into the field. In addition, a small-scale study of interobserver agreement under field conditions was conducted. All results indicated that satisfactory levels of agreement had been established and maintained (Stallings and Broussard, 1977).

A particularly sensitive issue having to do with observer effects arose early in Phase III, when late Phase II analyses suggested that there might exist systematic differences in perspective linked to the race of the observer. The existence of these effects could not be regarded as proven, because race of observer was partially confounded with the race of the child or caregiver under observation and with various center characteristics. Nevertheless, to guard against possible distortions due to race of observer, Phase III spring observation procedures were modified. According to the modified plan, every child was to be seen by two different observers, one black and one white. This modification was strongly urged by black consultants to the NDCS (Stallings and Broussard, 1977). Despite formidable difficulties of recruitment and scheduling, SRI came close to full implementation of the plan (Table 3.1). This procedure eliminated any confounding between policy variables and race of observer. Moreover, it made possible a much more precise estimate of the magnitude of observer effects than would otherwise have been possible. These estimates played an important role in broader investigations of the reliability of the study's observation measures.

SRI hired and trained 46 observers in both fall and spring. Each time nine observers conducted structured observations exclusively, while the remaining 37 conducted natural observations in classrooms. Between fall and spring, the number of observers who were members of minority groups was increased from 12 to 20, or 44 percent of the total. These particular observers completed 44 percent of all observations, close to the 50 percent ideally required by the study procedures. A minimum of 30 percent of observations in each center was conducted by minority observers. All observers were female. Distributions of age and education were fairly similar across sites; most observers were college graduates between 30-35 years of age.

Training of Observers*

Each observer trainee received a home training kit a few days before the observation training began, including a manual of code definitions, a detailed explanation of the coding procedures, and a sample coding sheet.

Observers who had not collected child focus observation data at any previous time attended training sessions for seven days. Experienced observers joined the

---

*This section summarized from Stallings and Broussard, 1977.

Table 3.1

CHILD FOCUS OBSERVER CHARACTERISTICS

Fall 1976

Classroom Observations	Number	Percent Minority	Average Age	Average Years of Education
Atlanta	19	32	35	17
Detroit	10	30	38	16
Seattle	8	12	29	17
Structured Observations				
Atlanta	5	40	32	17
Detroit	2	50	37	16
Seattle	2	0	31	17

Spring 1977

Classroom Observations	Number	Percent Minority	Average Age	Average Years of Education
Atlanta	20	50	36	16
Detroit	10	50	35	16
Seattle	7	29	33	16
Structured Observations				
Atlanta	5	40	35	16
Detroit	2	50	32	16
Seattle	2	0	31	16

trainers were paired in different combinations to form teams at the three sites.

At the conclusion of the training sessions, the performance of child focus trainees was assessed by means of a criterion videotape. The trainees recorded 115 examples of child behavior from the criterion videotape and 97 examples from the same videotape after they had been observing in the field for two weeks. Recordings of all observers who completed training were compared frame-by-frame with the criterion videotape. Those frames on the tape that were coded incorrectly by 50 percent or more of the observers were deleted from the data. After the bad frames were deleted, observer scores were computed as percent agreement with the criterion established by the SRI trainers. Forty-five of 51 trainees passed the original criterion videotape testing (passing score was 75% or higher).

In general, observers' scores improved somewhat from pretest to posttest. All of the 42 observers who participated in the retesting session scored 80 percent correct or better at that time, and mean accuracy was 93 percent. In addition, SRI conducted a field test of interrater agreement to address the issue of racial differences in coding patterns that had arisen in Phase II. Seventeen pairs of observers were formed, each with one black and one white member. Each pair coded the activities of the same child for one hour. Interobserver comparisons were possible for 45 activity codes, of which only three showed significant differences in overall frequency between black and white observers. Training of observers and results of various tests of observers' accuracy are described in more detail in SRI's Phase III report.

Reliability of Child Observation Codes

Effects analyses of the CFI were aimed primarily at determining to what extent variations in child activities and behavior are related to variations in group size, staff/child ratio and caregiver qualifications. A determination of the reliability of the CFI codes was an important step in the development of analytic strategy. First, estimates of reliability helped in setting the proper unit of analysis by identifying the level of aggregation--person (child or caregiver), class or center--for which the data are most reliable. Second, they helped establish the mathematical limits of the analyses to be performed--the degree of statistical power to detect relationships and the degree of bias likely to be present in estimating the strengths of relationships. When reliability coefficients are modest, meaningful analyses can nevertheless be conducted if the sample provides enough degrees of freedom. However, under such circumstances, genuine but small relationships may not reach conventional levels of statistical significance.

While most researchers who use observation-based measures report only "inter-rater reliabilities" (usually percentages of agreement or correlations between scores generated by pairs of observers) or stabilities of measures across occasions (day-to-day correlations), other mathematical techniques have been developed to a point of considerable sophistication. The essential ideas were first elaborated by Medley and Mitzel (1963) and have been most fully elaborated by Cronbach, et al., (1972).

The reliability of the CFI measures was assessed in three ways: through observer agreement (as a part of SRI training described above), through examination of the stability of the measures across timepoints, and through components of variance analysis. In general, the reliability analyses indicated that the measures were sufficiently reliable to support the effects analyses at a classroom level, but not at the child level. With all observations for a group aggregated, a significant proportion of the variance in the measures will be expected to be systematic and potentially explainable by the policy measures. On the other hand, the analysis indicated that the key NDCS variables could be expected to account for only moderate proportions of the total variation.

Stability from fall to spring and from free play to adult-directed observation periods was assessed by computing correlations between frequencies of the same code in day care groups at different times (Table 3.2). The overall pattern of correlations suggests that codes were more stable in the fall than in the spring. This could be due to a clearer differentiation of activities by springtime (when the group had been together for six months) or to the fact that children in each classroom were generally seen by a single observer in the fall but at least two observers in the spring. As might be expected, stability from fall to sring was lower than stability across observation types within a data collection period. Most relationships were statistically significant at a moderate level, however, suggesting that underlying stability is found for these codes. The major proportion of variation in child behavior may be associated with factors other than the NDCS policy measures (for example, child age and fmaily background).

Neither interobserver agreement nor stability over time conveys all of the information needed to judge the suitability of CFI measures for use in effects analyses. What matters is how well the measures characterize children, classes or centers when appropriately averaged. In more technical terms, the critical element is the relative

Table 3.2

CHILD FOCUS OBSERVATION CODE STATILITY: OBSERVATION TYPE AND DATA COLLECTION PERIOD

	Free Play/Teacher Directed Activities		Fall 1976/Spring 1977	
	Fall 1976 N=117	Spring 1977 N=116	Free Play N=114	Teacher-Directed Activity N=114
A. Child Activity Codes				
Group structured activity	.27**	.18	.12	.16
Group open activity	.41**	.19	.14	.20*
Monitors environment	.38**	.22*	.21*	.18
Gives opinions	.62**	.32**	.18	.12
Wanders, does nothing	.40**	.51**	.31**	.26*
Group passive activity	.18	.19	.16	.12
Moves with purpose	.33**	.20:	.15	.16
Individual open activity	.25*	.14	.13	.23*
Adds prop or idea	.30**	.16	.17	.18
Considers, contemplates problem	.34**	.38**	.26*	.20*
Individual structured activity	.21*	.18	.11	.22*
Receives general comments	.44**	.30**	.22*	.18
Receives information, guidance	.28	.35**	.17	.11
Receives demands, requests	.51**	.37**	.23*	.21*
B. Object Codes				
Attention to adults	.43**	.36**	.43**	.08
Attention to children	.60**	.35**	.27**	.30*
Attention to groups	.24*	.27**	.38**	.27**
Attention to environment	.45**	.25*	.21*	.16
C. Activity Continuity Codes				
Longest activity	.07	.06	.08	.10
Not involved in task	.50**	.34**	.44**	.11

*p<.05
**p<.01

importance of different sources of variance for each of the CFI measures. The more variance that is linked to other factors, including but not limited to observer differences and fluctuation in caregiver behavior, the less dependable is the measure as a descriptor of the chosen unit.

To assess the relative importance of the major potential determinants of CFI scores, a variance components analysis was performed (Singer, Affholter and Goodrich, 1978). That effort revealed that for all of the CFI codes analyzed, the occasion of observation was the major determinant of variation, followed by class and observer. Variation at the child level contributed little to the total. However, since so many observations were conducted in each class, the group agregate values are considerably more reliable. Variance associated with observers is a relatively large proportion of total variance, although the influence of observer variance upon group aggregates is reduced in the spring observation due to the use of multiple observers in each classroom. Reliabilities of selected CFI codes at the child and class level are shown in Table 3.3. The table suggests that reliabilities of child behavior were extremely low at the child level; the codes did not appear to describe enduring traits or stable behavior of children. However, the measures showed class level generalizabilities that were adequate for effects analyses, given the degrees of freedom involved. Class level generalizabilities ranged from .14 to .60, clustering in the neighborhood of .35.

These results shaped the choice of units of analysis for the CFI. In addition, they provided a context for interpreting quantitative findings. The results suggested that the overall explanatory power of regression models would be limited. Even if very strong underlying relationships between the policy variables and dependent variables were to exist, generalizability limitations would restrict the explanatory power of CFI regression models such that even R^2's of .4 would be difficult to obtain. The larger implication was that relatively modest relationships should be taken seriously. The NDCS was a search for signals in a noisy environment; a signal loud enough to detect was likely to be stronger than it seemed against the background noise.

Table 3.3

GENERALIZABILITIES OF CHILD FOCUS OBSERVATION CODES

	Child	Class
A. Child Activity Codes		
Group structured activity	.05	.42
Group open activity	.03	.37
Monitors environment	.02	.21
Gives opinions	.07	.52
Wanders, does nothing	.06	.60
Group passive activity	.00	.14
Moves with purpose	.05	.46
Individual open activity	.01	.18
Adds prop or idea	.04	.44
Considers, contemplates problem	.03	.32
Individual structured activity	.00	.14
Receives general comments	.03	.41
Receives information, guidance	.06	.52
Receives demands, requests	.03	.31
B. Object Codes		
Attention to adults	.03	.34
Attention to children	.02	.31
Attention to groups	.03	.35
Attention to environment	.03	.40
C. Activity Continuity Codes		
Longest activity	.01	.18
Not involved in task	.02	.27

CHAPTER FOUR: SELECTION AND CONSTRUCTION OF DEPENDENT MEASURES

The major portion of the effects analysis focused upon fifteen measures of child behavior. Eleven of these measures are frequently-occurring individual CFI codes. Four are the codes denoting the object of the child's attention--orientation to adults, to individual children, to groups of children and to the environment--which are emphasized because they describe the child's global interaction patterns. The seven other individual codes included Verbal Initiative, Noninvolvement, Aimless Wandering, Task Persistence, Passive Group Time, Environment Monitoring, and Purposeful Movement. The four remaining dependent measures--Reflection/Innovation, Cooperation/Compliance, Classroom Activity Balance, and Interest/Participation--were multicode measures developed for the analyses. Definitions of the measures are described below, including the process of code construction. The general measurement strategy was to describe behavior in the day care center as comprehensively and objectively as possible, in terms of fine-grained codes. Data were then reduced by combining frequencies of codes that were conceptually related and empirically correlated. Efforts were made to create summary variables that bore some relationship to constructs previously used in the developmental literature, but primary weight was placed on empirical patterns evident in the data.

To choose appropriate combinations of codes, frequencies and intercorrelations of the various codes were examined, at all levels of aggregation--child, class and center. Data were also examined separately for fall and spring, for the Atlanta Public School classrooms, and for the three sites of the 49-center study.

In addition, principal components analyses were performed on child- and class-level data from the fall and spring samples, in an independent effort to reduce the set of codes to a few summary dimensions. The principal components analysis proved unrevealing. The resulting dimensions accounted for little variance and were not readily interpretable. Nor were they especially stable from fall to spring. Moreover, some "dimensions" were dominated by one or two particularly frequent codes. Consequently, conceptual coherence and simple correlations among codes were the primary bases for deciding how to combine codes to form broader constructs. Where individual codes were sufficiently frequent, important and distinctive in meaning from other codes, they were treated as variables in themselves. (Again, as indicated earlier, some infrequent codes representing psychologically important events were treated differently and are discussed separately.)

Along with definitions of the measures that follow, each section below contains information on the consistency of the measure across adult-directed and free play activity periods (indicating the degree to which the measures characterize classrooms rather than activity segments within classrooms). Differences in behavior frequency in children of different ages are reported, and selected correlations among the measures are reported wherever these help clarify the meaning of a particular measure. Finally, stabilities of measures from fall 1976 to spring 1977 are also reported. Stability correlations identify those constructs for which center classrooms retain their relative frequency rankings from fall to spring, as opposed to those constructs for which classrooms shift noticeably in relative frequency ranks. These measures give some indication of which behavior patterns are established rapidly during the day care year* and which patterns take shape gradually from fall to spring. However, the correlations are somewhat underestimated because of changes in observation procedures from fall to spring discussed in Chapter Two and because of shifts of enrollment within classes.**

Reflection/Innovation

Two codes--CONSIDERS, CONTEMPLATES OR TINKERS, and ADDS PROP OR IDEA came closest among all CFI codes to capturing thoughtful, creative problem-solving behavior on the part of children. Because of their low individual frequencies and positive correlations (.34 in fall, .30 in spring), the two were summed to form a statistically more robust variable, REFLECTION/INNOVATION. Frequencies of the construct tended to be consistent across activity periods ($r=.42$, $p<.01$ in fall; $r=.37$, $p<.01$ in spring) but were unstable from fall to spring.

Verbal Initiative

The single code GIVES OPINIONS, REFERENCES, INFORMATION, COMMENTS was treated as a separate variable indicating the degree of verbal self-assertiveness exhibited by children and expected or accepted by caregivers. Frequencies of VERBAL INITIATIVE were consistent across activity types (r=.62, p<.01 in fall; r=.32, p<.01 in spring) but had only modest fall-to-spring stability (r=.18, p<.05 for free play; r=.12, n.s. for adult-directed activity).

---

* The "day care year" is not as sharply defined as the school year, with a clear beginning in fall and in spring. However, formal and anecdotal NDCS data from both the Supply Study and main cost-effects study show that there is a major influx of new children in the fall, accompanied by an exodus of children who have reached school age. There is also a drop off of enrollment during the summer months.

**Correlations of code frequencies between free play and teacher-directed activities are based on 117 classrooms in fall and 116 in spring. Fall-spring stability correlations are based on 114 classrooms that existed at both time points, although shifts in enrollment occurred within those classrooms.

Cooperation/Compliance

Seven of the RECEIVES codes are accompanied by supplementary codes indicating whether the child's response is appropriate. The seven relevant categories of action or statement directed toward the child are (1) general comments, (2) information or guidance, (3) requests to play or share, (4) demands or requests other than requests to play or share, (5) rules or corrections, (6) punishment or threats, and (7) praise. Percentages of appropriate responses, shown in Table 4.1, ranged from 48 percent for punishment and threats to 87 percent for comments, information and guidance. An index of COOPERATION/COMPLIANCE was computed as the ratio of all active appropriate responses to all instances of these seven RECEIVES codes. In the fall, older children showed higher frequencies of COOPERATION/COMPLIANCE than younger children (p<.05), but no age differences were evident in spring--perhaps indicating a progressive socializing effect for younger children. Cooperation was at best marginally consistent across activity periods (r=.18, p<.05 in fall; r=.08, n.s., in spring). Cooperation during free play was moderately stable from fall to spring (r=.25, p<.01) but cooperation during adult-directed activity was not stable (r=.06, n.s.).

Noninvolvement (No Task)

The degree to which children are uninvolved in classroom activities is directly recorded by the activity continuity code NO TASK ("Task" is broadly defined and includes play and exploration as well as teacher-assigned activities). This index of noninvolvement was consistent across activity types (r=.50, p<.01 in fall; r=.34, p<.01 in spring) and was stable from fall to spring for adult-directed activity (r=.44, p<.01), but much less so for free play (r=.11, n.s.).

Aimless Wandering

Like noninvolvement, aimless wandering--measured by the frequency of the code DOES NOTHING, WANDERS--is an index of the degree to which children are not engaged in classroom activities. The two variables are correlated (r=.28, p<.01, for free play, and r=.45, p<.01, for teacher-directed activity). However, the two were not summed to form a single construct because they were incommensurate. DOES NOTHING, WANDERS was an activity code, one of 50 possible, whereas NO TASK was a continuity code, one of three possible. NO TASK was often recorded along with DOES NOTHING, WANDERS, accounting in part for their correlation and rendering their sum meaningless. The frequency of AIMLESS WANDERING was consistent across activity types (r=.40, p<.01 in the fall, and r=.52, p<.01, in the spring) and was moderately stable from fall to spring (r=.28, p<.01 for all activity types pooled).

Task Persistence

The concepts "task persistence" and "attention span" commonly refer to a child's tendency or ability to devote sustained effort to a single pursuit. Increasing the young child's capacity in this area is often regarded as an important goal of early education. The focus here is less on task persistence and attention span as individual traits than on closely related characteristics of the classroom, namely demands made and opportunities provided for sustained activity. The CFI provides an indirect measure of these constructs. The activity continuity code designated OLD ACTIVITY marks every occasion on which a child continues an activity from one twelve-second interval to the next. By summing durations of all intervals so marked, between the outset of the activity (indicated by a NEW ACTIVITY code) and its termination (indicated by another NEW ACTIVITY code or a NO ACTIVITY code) it is possible to measure the total duration of every activity in the twenty-minute observation period to the nearest twelve seconds. The mean duration of each child's longest activity is approximately 11 minutes. Phase III data, consistent with Phase II findings and previous research, show that activities last longer, on the average, in groups of older children than in younger groups. Moreover, activities last longer in groups where structured activities predominate. The correlation between activity length and the classroom activity balance was -.37 (p<.01) in fall and -.48 (p<.01) in spring. However, LONGEST ACTIVITY was neither strikingly consistent across activity types nor stable from fall to spring.

Patterns of Interaction

The interaction variables are based directly on the object codes of the CFI. They describe the distribution of children's attention to the social and physical environment at a relatively coarse level, distinguishing the amount of attention directed to adults, to other children, to groups of children, and to the physical environment.

Orientation to Adults was, predictably, twice as frequent in caregiver-directed activity as in free play. However, frequencies showed fairly high correlations across the two types of activity period (r=.43, p<.01 in fall, r=.36, p<.01 in spring), indicating that some groups of children were consistently more adult-centered than others, regardless of prevailing activities. The construct was more stable from fall to spring for free play (r=.43, p<.01) than for adult-directed activity periods (r=.08, n.s.).

Orientation to Individual Children also showed substantial correlations between free play and teacher-directed activities (r=.60, p<.01 in fall; r=.35, p<.01 in spring), again indicating a consistent focus of some classrooms on child-child interchange. Combined frequencies of this variable across the two types of activity period were moderately stable from fall to spring (r=.29, p<.01). However, the Atlanta Public Schools subsample, which showed a particularly high level of orientation to children in the fall, also showed a drop from fall to spring which was not observed in any of the 49-center study sites.

Orientation to Groups was included as a dependent measure primarily to determine whether children's contact with their peers affected by classroom composition, specifically whether their attention is directed to group rather than solitary or one-to-one activity as total class size grows. Fall-to-spring correlations for this measure were .38 (p<.01) for free play and .27 (p<.01) for teacher-directed activity. Consistency across teacher-directed and free play activities was .24 (p<.05) in fall and .27 (p<.01) in spring.

Orientation to the Environment also showed consistency across activity types (r = .45, p<.01 in fall; r=.26, p<.01 in spring). However the variable showed only modest stability across data collection periods. (The fall-spring correlation for the combined score across activity periods was .18, p<.01). In addition, there were noticeable differences across sites in patterns of change from fall to spring. In both fall and spring there was a marginally significant tendency for younger children to show more attention to the environment and less to adults and peers than did older children (p=.11 in fall, p=.06 in spring).

Interest/Participation. A global variable was constructed reflecting the degree to which children in a class are actively involved in its social and educational activities. INTEREST/PARTICIPATION was computed as the sum of many codes (GROUP AND INDIVIDUAL OPEN, EXPRESSIVE ACTIVITY; CONSIDERS, CONTEMPLATES OR TINKERS: ADDS PROP OR IDEA: ACTS CREATIVELY OR SOLVES PROBLEM: OFFERS TO HELP OR SHARE; DEFENDS RIGHTS: MOVES WITH PURPOSE; SELECTS ACTIVITY (alone or with others), ASKS FOR INFORMATION; ASKS PERMISSION TO SHARE; GIVES OPINIONS; ASKS FOR RECOGNITION; GIVES ORDERS OR DIRECTS OTHERS; INTRUDES PLAYFULLY). The construct is related to a behavior cluster that has emerged repeatedly in studies of preschool children in group care settings and that is associated with children's later social adjustment and cognitive achievement.[1] A similar construct also emerged during Phase II of the NDCS.[2] In both Phase II and Phase III, codes comprising the construct were positively correlated with each other and negatively correlated with codes indicating noninvolvement. INTEREST/PARTICIPATION also was positively related to task persistence (r=.22, p<.05 in fall; r=.26, p<.01 in spring). NON-INVOLVEMENT showed negative correlations in the .3-.4 range with INTEREST/PARTICIPATION in both free play and teacher-directed activity periods. Thus, NON-INVOLVEMENT and INTEREST/PARTICIPATION together tend to array classrooms along a general dimension indicating the degree to which children are integrated into classroom activities. A similar bipolar dimension emerged in the studies as well as in Phase II of the NDCS. In spring, high levels of COOPERATION/COMPLIANCE tended to accompany high levels of INTEREST/PARTICIPATION and low levels of NON-INVOLVEMENT. (No significant relations were found in fall.) In short, though the relevant correlations were not strong, INTEREST/PARTICIPATION was part of a broad cluster of positive dynamics in the classroom.

Classroom Activity Balance

A variable was constructed to characterize the global dynamics of the day care group activities. The most commonly used CFI codes were PARTICIPATES IN GROUP ACTIVITY--CLOSED, STRUCTURED and PARTICIPATES IN GROUP ACTIVITY--OPEN-ENDED, EXPRESSIVE. These two codes represented about one-third of all activities recorded. When individual structured and open-ended activities are pooled with the respective group activity codes, all four together accounted for over 37 percent of the codes recorded. Class-level correlations between frequencies of structured and open-ended activities were negative and substantial in both the fall ($r=-.36$, $p<.01$) and spring ($r=-.63$, $p<.01$), indicating that classrooms tend to be characterized by one type of activity or the other.

It should be noted that activities defined as "closed, structured" are not necessarily equated with educational activities. Rather, the codes represent activities that have a clearcut end point or achievable goal, whereas open-ended expressive activities do not. Either type of activity can be educationally or developmentally valuable. Nevertheless the two types of activity codes seem to capture distinctive classroom styles.

The CLASSROOM ACTIVITY BALANCE, designed to locate a given classroom on the structured/open-ended dimension, is constructed BY subtracting the sum of frequencies of group and individual structured activities from the sum of frequencies of group and individual open-ended activities. This difference score averaged -.06 in the fall and -.04 in the spring, indicating a slight prevalence of structured over open-ended activity, and very little change with time in the overall balance among Phase III centers. The relative ranking of different classrooms on the unstructured/open-ended dimensions was moderately stable from fall to spring ($r=.36$, p<.01). Open-ended activities were more prevalent in classes with younger children.

Activity Transition

Three frequently occurring codes seemed to involve, at least partially, a form of transition from one activity to another. The single codes for MONITORING ENVIRONMENT, MOVES WITH PURPOSE, and GROUP PASSIVE ACTIVITY were considered independently as variables. Each of these codes appeared to suffer from multiple meanings requiring a more elaborate description of the context than was possible for the CFI. For example, group passive activity included waiting in line and television watching as well as storytime.

MONITORS ENVIRONMENT was consistent across activity types (r=.61, p<.01 for fall and r=.37, p<.01 for spring) but was only moderately stable from fall to spring (r=.07, n.s. for free play, r=.21, p<.01 for adult directed activity). MOVES WITH PURPOSE was also relatively stable over activity types (r=.40, p<.01 for fall and r=.22, p<.01 for spring) but only moderately stable over data collection periods (r=.10, n.s. for free play; r=.28, p<.01 for adult directed activity). GROUP PASSIVE ACTIVITY provided little evidence of stability across either activity types (r=.01, n.s. for fall and r=-.02, n.s. for spring) or data collection periods (4=.05, n.s. for free play and r=-.01, n.s. for adult directed activity).

CHAPTER FIVE: SELECTION OF INDEPENDENT MEASURES

The selection of independent and dependent variables was motivated by a decision made at the outset of the study to focus attention on those aspects of the quality of day care that bear directly on the child. In effect, ACYF (and its contractors) took the position that the primary goal of day care purchasing standards is to ensure the best possible environment for the most children. Other goals of day care--e.g., freeing parents to work, serving as a vehicle for delivery of social services to parents, employing low-income people as staff and fostering their development as professionals--were recognized as legitimate and important but were not central to the study.

As a consequence, in selecting regulatable center characteristics for intensive investigation as independent variables, priority was given to those deemed most likely to affect children's daily experiences, namely the composition of the classroom (principally group size and staff/child ratio) and the qualifications of caregivers (education and experience). Other center characteristics (space, equipment and materials; center philosophy and curriculum; director qualifications; stability of caregiver/ child relationships; availability of nutrition and health services; availability of other supplementary services and specialists; opportunities for parent involvement) were examined in descriptive and exploratory fashion to determine whether any appeared to have major effects on classroom processes and child outcomes. However, in light of preliminary results which suggested that most of these variables had minimal effects on the particular outcome measures chosen, only a few of the variables were investigated further, and then only to a limited extent.

Two types of independent variables were tested in the effects analyses: *background* variables, such as age, sex and race of children as well as socioeconomic characteristics of families and of the community served by the particular center, and *policy* variables, i.e., center characteristics subject to regulatory control. While background variables are unregulatable and therefore not of direct policy relevance, their effects had to be taken into account in assessing the effects of the policy variables.

Information on background characteristics of children and their families was gathered through interviews with parents. Background information included family income, sources of income, parents' education and occupation, length of parents' employment, number of siblings and number of adults living in the house. Age, sex and race of children were verified. In addition, census data were used to provide background information on demographic characteristics of the community, chiefly its socioeconomic and racial composition.

A set of variables that describe the socioeconomic characteristics of the children's families was aggregated from the child level to the class level as potential independent measures. These included mean family income, mean mother's education, mean proportion of single parents, mean number of children under 12 in the home, and mean proportion of black children in the class. The simple correlations of these demographic variables with the dependent measures indicated few significant relationships, more often associated with the fall data collection than with the spring. Since the five demographic variables were strongly intercorrelated, (Table 5.1) and because was five SES covariables was a large number to enter into the regression analyses as independent measures, a clustering was completed to represent the average socioeconomic status of the children in each class.

Table 5.1

INTERCORRELATION OF SOCIOECONOMIC MEASURES IN NDCS CLASSROOMS

	Mother's Education	Family Income	Proportion of Single Parents	Number of Children Under 12	Proporation of Black Children
Mother's Education	—	.72	-.55	-.43	-.25
Family Income	.71	xx	-.66	-.42	-.43
Proportion of Single Parents	-.51	-.70	—	.36	.39
Number of Children Under 12	-.47	-.39	.35	—	.31
Proportion of Black Children	-.26	-.41	.38	.32	xx

All values shown are $p<.05$; values above the diagonal are for fall (N=117), values below the diagonal are for spring (N=116).

The five variables were factor analyzed and a principal component factor score was assigned to each class. This principal component score typically fell between -1 and 1, with positive scores indicating higher SES, was the main SES covariable in the regression equations. In addition, caregiver race was used as a covariable.

The third covariable that was entered in regression computations was child age. The grouping of children by age is a well-established practice in day care classrooms. The 1968 federal regulations require slightly more staff per child and smaller numbers of children in groups of children three years of age as opposed to groups of four-year old children. Thus, if age where not controlled statistically, there might be a tendency for the less mature behavior of three-year old children to be associated with high staff/child ratio and small groups. This would not, however, be an accurate indication of the impact of group composition upon child behavior. Therefore, in all regression equations for the CFI variables, the average age of the children in the group was included as a covariable.

The major policy variables examined in the NDCS fell into two categories--those relating to classroom composition and those relating to caregiver qualifications. Three classroom composition variables were considered:

- NUMBER OF CAREGIVERS, defined as the total number of caregivers present in or assigned to a classroom;

- GROUP SIZE, defined as the total number of children present in or assigned to a class or to a principally responsible caregiver;* and
- STAFF/CHILD RATIO, defined as the number of caregivers divided by group size.

Caregiver qualifications variables consisted of total years of formal education (YEARS EDUCATION), presence or absence of childrelated education/training (SPECIALIZATION), day care experience--both prior to current job (PREVIOUS DAY CARE EXPERIENCE), and time in current center (CENTER EXPERIENCE).

Information on caregiver qualifications was gathered from interviews with caregivers. Information on variables related to classroom composition was gathered by two methods, one based on schedule or roster data and the other on direct observation, i.e., as part of the CFI. Schedule-based and observation-based measures of classroom composition were not always in close agreement (see Bache, 1980a). Differences between the two were primarily attributable to two phenomena--absenteeism and merging of classes. Because observations capture the group configurations actually experienced by the child and because they automatically take account of absenteeism and merging, observation-based measures were used in all the analyses reported in this paper.* Table 5.2 shows correlations among the major policy variables across the Phase III classrooms. The table indicates that the classroom composition variables are essentially uncorrelated with the caregiver qualifications variables, so that their effects can be easily separated. Within the cluster of qualifications variables, modest correlations exist--high enough to warrant caution in interpreting individual effects but not high enough to preclude identification of the most powerful variable(s). A similar correlation exists between GROUP SIZE and STAFF/CHILD RATIO for the spring observations.

---

*In all but a few NDCS centers, groups of children were assigned to particular rooms, supervised by a single caregiver or several caregivers. In a few "open classroom" centers, however, very large numbers of children (approaching 100 in extreme cases) were present in a single large room. Even in such centers, children clustered around individual caregivers or small teams dispersed around the room, though children were often free to move from group to group. Numbers of children in these smaller groups constituted the group size used for NDCS analytic purposes. Similarly, numbers of caregivers reflected the number of adults in physically separated groups.

Table 5.3 shows relationships among the policy variables and the set of background variables describing the children, families and communities served by the NDCS centers. Again, most correlations are small, indicating that effects of policy variables can be separated from those of particular background factors. Some moderate correlations do exist, however. Perhaps most important are the associations of staff/child ratio and staff qualifications in current centers with various indices of socioeconomic status: high ratios and experienced staff with some child-related education/training are found in centers serving low-SES families. This pattern of associations is tied to federal funding. Low-income children are served in federally funded centers, which are subject to higher FIDCR ratio and training requirements and

---

*Three distinct sets of observation data were collected. One set was collected on a regular basis by NDCS staff employed full time at each center during Phases II and III; this set was used in analyses of non-behavioral data. A second set of counts was made in conjunction with behavioral observations of caregivers, and a third in conjunction with observations of children; these counts were used in the corresponding behavioral analyses. Behavioral observations are described below and in later chapters.

Table 5.2

CORRELATIONS AMONG THE MAJOR POLICY VARIABLES

	Group Size	Staff/ Child Ratio	Years of Education	Child-Related Education/ Training	Previous Day Care Experience	Experience in Current Center
Classroom Composition						
Group Size	--			-.23		
Staff/Child Ratio	-30	--				
Staff Qualifications						
Years of Education			--	.20		
Child-Related Education/Training			.36	--	.28	
Previous Day Care Experience				.28	--	
Experience in Current Center						--

All values shown are p<.05; values above the diagonal are for fall (N=117); values below the diagonal are for spring (N=116).

Table 5.3

CORRELATIONS AMONG POLICY AND BACKGROUND VARIABLES

	Fall, 1976 (N=117)		Spring, 1977 (N=116)	
	SES of Family	Age of Children	SES of Family	Age of Children
Classroom Composition				
Group Size		.21		.22
Staff/Child Ratio	-.27	-.24	-.22	-.20
Staff Qualifications				
Years of Education				
Child-Related Education/Training	-.28		-.34	
Previous Day Care Experience			-.33	
Experience in Current Center	-.34			

All values shown are p<.05.

which pay slightly higher wages and have lower staff turnover rates than do parent-fee centers. This pattern of relationships reaffirms the point that background factors, such as socioeconomic status, must be controlled when exploring relationships between policy variables and various measures of children's development.

CHAPTER SIX: RESULTS OF CFI ANALYSES

The CFI analyses are reported in four groups. First, the effects of staff/child ratio were tested in a quasi-experiment with a three-group ANOVA design. The results for the quasi-experiment suggested that the manipulation of ratio did not result in systematically different child activities or behavior. Second, regression analyses were performed for the commonly used codes and combinations. Third, rare but important behaviors were analyzed by logit techniques. Finally, regression analysis was used to investigate the behavior of children in a structured situation.

The variables examined in these analyses are combinations of codes that utilize virtually all of the coded child behavior. The CFI gathered a broad range of information relevant to many different behaviors of children. Day-to-day behaviors were observed in considerable detail. The resulting measures should not be used directly or individually to form judgements about individuals or the dynamics of different classrooms. Rather, it is important to identify patterns of child activity that, taken as a whole, could reasonably be judged as beneficial to children.

There are three reasons for this approach. First, the reliabilities of individual codes and measures, while acceptable for analytic purposes, are not high enough to support the use of a single measure. Second, there are numerous significant intercorrelations among the observation measures, indicating that the measures are not independent. Third, the CFI measures do not represent behaviors that have a positive or negative meaning independent of the relative level of the behavior in the entire data set. While an "optimal" level for each behavior may exist, that value has not yet been determined. An example of this is the amount of time that children are uninvolved in meaningful activity. A preschool environment where children were continuously involved in some activity with no time during the day to be by themselves might be regarded as an over-stimulating or over-regimented day care group. Nevertheless, a relatively high level of inactivity may be seen as an essentially negative index when it is related to independent measures in similar ways as frequent wandering, low levels of verbal initiative and/or cooperation.

Experimental Study of Ratio Effects

As described earlier, the effects of staff/child ratio were tested in a quasi-experiment in 49 of the centers. The question tested was whether children in high-ratio classrooms with an experimentally induced increase in ratio behaved differently when compared with children in either untreated low-ratio or untreated high-ratio classrooms. In a series of one-way analyses of variance comparing child behavior in these three types of situations, there were few detectable effects of ratio. A significant result ($p<.05$) occurred in only two cases. During adult-directed activities, children in the high* ratio groups wandered less and exhibited more task persistence. Since two significant results from the 30 analyses conducted could have occurred by chance, little confidence can be placed in these results as impacts of the NDCS manipulation.

Similarly, in the APS experiment, few significant effects could be found for the manipulated variables (STAFF/CHILD RATIO and LEVEL OF EDUCATION). Children in high ratio APS classrooms cooperated more readily during both free play and adult-directed activities, and oriented to adults more frequently during free play. Children in groups with a high level of education appeared to cooperate more readily during adult-directed activities. Although these results suggest a mildly positive impact for the research variables, there is not sufficient strength nor consistency in these outcomes to conclude that the NDCS manipulations had a major effect upon the activities and behavior of children.

---

*Staff/child ratio, defined as the number of caregivers divided by group size is higher in situations where relatively more adults are present.

Regression Analysis

The CFI regression model included six policy variables and two covariables. The six policy variables were observed GROUP SIZE, observed STAFF/CHILD RATIO, caregiver YEARS OF EDUCATION, education in a child-related field (SPECIALIZATION), experience in day care prior to employment at current center (PREVIOUS EXPERIENCE), and experience in current center (CENTER EXPERIENCE). All measures, dependent and independent, were averaged to the classroom level. Thus, the measures of caregiver qualifications represent averages for the staff in the classroom. The two covariables entered were average age of children in the class (CLASS AGE) and a class-level measure of socioeconomic status (FAMILY SES).

For the regression analysis of CFI data, both standard regression and biweighted regression equations were produced. Biweighted regressions were used because ordinary least square estimators are neither robust nor resistant where data are not normally distributed (Mosteller and Tukey, 1977). Resistance concerns the sensitivity of obtained estimates to extreme values of a small number of sample observations. Robustness refers to the relative stability in the population from which the sample is drawn. In cases like the NDCS, where class level regression models are estimated on sample and subsamples ranging in size from 28 to 117, the relative nonresistance and nonrobustness of ordinary least squares may be problematic. Outlier values may suppress or elevate effects which are not applicable for the entire sample.* Biweighting is one way to approach the problem of estimation under such conditions.

Summary and Discussion of Regression Analyses. A summary of the significant regression results (Table 6.1) suggests a meaningful profile of child activity and behavior that is related to both the day care year and key NDCS policy measures. A discussion of the analysis of each individual measure follows this review of the overall results for covariables and policy measures.

Covariables. In the fall, both CHILD AGE and FAMILY SES frequently related to the behavior of preschool children in day care groups. Older children offered more opinions, were more cooperative, and were oriented towards other children more frequently than were younger children. Transitional activity in older day care groups was more often associated with movement than with environmental monitoring. By spring, several of the fall trends had been softened. Age trends for VERBAL INITIATIVE, COOPERATION and TRANSITIONAL ACTIVITY were no longer significant. Age differences in PATTERNS OF INTERACTION persisted (older children spent relatively more time oriented to other children and groups and less time oriented to adults). Perhaps in preparation for entry into school there was a tendency for more formal, structured activity in groups of older children by springtime.

---

*Outliers may take many forms; they may be deviations from the regression surface, extreme values of the observed values relative to the rest of the sample, or both. Biweighted regression essentially attempts to deal with deviations from the regression surface that are large. Other "outliers" may "unduly" influence the regression fit without necessarily being extreme deviations from the fitted surface. Such observations are said to have high "leverage," and another set of diagnostic tools is required for examining data with potential problems from a few observations with high leverage (see Hoaglin and Welsch, 1978).

Table 6.1

SUMMARY OF SIGNIFICANT REGRESSION RESULTS

															Patterns of Interaction								Transition Codes				
	Reflection/ Innovation		Verbal Initiative		Cooperation		Non-Involvement		Wandering		Task Persistence		Interest/ Participation	Classroom Activity Balance	Orientation to Adults		Orientation to Children		Orientation to Groups		Orientation to Environment		Monitors Environment		Moves with Purpose		Group Passive Activity
Fall Observation Period Covariables	FP	AD	FP	AD	FP	AD	FP	AD	FP	AD	FP	AD	Comb	Comb	FP	AD	FP	AD	FP	AD	FP	AD	FP	AD	FP	AD	AD
Age of Children			+	+[a]		+																					
SES of Family	+				(-)		+	+			-				-		+	+					-		(+)		
														+							-	-			(-)		+
Group Composition																											
Group Size		-	-[a]	-																							
Staff/Child Ratio												+	-		-	(-)				(+)			+	+	+	+	
										-		+		-									(+)			(+)	
Caregiver Qualifications																											
Specialization							-[a]	-[a]																			
Years of Education	+		(-)	-	-				-	(-)	+	+		(-)				(-)									
Previous Day Care Experience											-	(-)		(+)											+		
Experience in Current Center																											
														+										-			
Spring Observation Period Covariables																											
Age of Children																											
SES of Family							+	(+)			-	(-)		-	(-)	(-)		+		+							
														+		(+)											
Group Composition																											
Group Size	(-)	(-)	-	(-)	-[a]	(-)	+[a]		+	(+)		(+)															
Staff/Child Ratio								(-)	(-)	-	(+)	+			-	-			+[a]	+[a]			(+)	(+)			(+)
Caregiver Qualifications																											
Specialization		+			+		-				+	+	(+)														
Years of Education	(+)				-						-	(-)															
Previous Day Care Experience											+							+							(+)	(+)	
Experience in Current Center																					(-)	(-)					
	+																										
																											-

FP: Free Play Observation Periods; AD: Adult Directed Observation Periods

†Results noted were significant at p<.05; results significant at .05<p<.15 shown in parentheses.

[a] Results significant at p<.15 in both APs and 49-center samples.

Similarly, differences related to FAMILY SES were weaker in the spring than in the fall, suggesting a homogenizing effect of the group care experience. During the fall observations, children from higher SES families were somewhat less cooperative, less involved and less persistent at activities, but they engaged more often in reflection/innovation. Those same children were also found more often in open, unstructured activities and passive group activities. Several of these trends persisted in the spring data base. Higher SES children continued to participate relatively more often in unstructured activities where involvement and persistence at a task were less. However, fall observation differences concerning REFLECTION/INNOVATION, COOPERATION and TRANSITIONAL ACTIVITY did not reappear in the spring. This pattern suggests that NDCS centers serving relatively higher SES children were less structured in their approach to activities and perhaps took a longer time to organize the children into functioning groups while leaving time for individual activities and thought during the period when many children were entering or re-entering the center on a full-time basis.

Group Composition. Relationships of child behavior to the size of the day care group were strong and consistent during both fall and spring observation periods and were frequently significant both in simple correlations and in regression equations. The strength of the relationship of group size to child activity was persistent across both free play and teacher-directed activities and across the day care year. Examination of biweighted regression results indicates that outlier groups did not greatly influence the GROUP SIZE result and actually strengthened those results in the fall observation data base.

In both fall and spring, children in smaller groups offered opinions more frequently and engaged in more reflection/innovation. Small groups of children also were more often oriented to adults and less often to groups of children than large groups. In addition to relationships that were persistent over the day care year, several other types of child behavior were related to group size in the spring but not in the fall, suggesting the emergence of some relatively undesirable trends in large groups. In the spring data base, large groups of children cooperated less readily and had more activity involvement and less wandering. Also, large groups remained in activities longer and, perhaps a related result, were found more often in passive activities.

In contrast, STAFF/CHILD RATIO was infrequently related to child behavior during either data collection. However, higher ratios were associated with less wandering and with greater task persistence both in the fall and in the spring. In addition, higher ratios were associated with more unstructured activities and more transitional activity (environmental monitoring during free play; moving with purpose during adult-directed activity) during the fall. Although high ratios appear to have less pervasive effects than small groups, the observed relationships suggest a moderately positive influence for higher ratios.

Caregiver Qualifications. None of the caregiver qualifications had especially powerful or pervasive effects. However, classrooms with a relatively high proportion of staff having child-related preparation (SPECIALIZATION) were marked by higher child involvement and greater task persistence at both data collection points. Also, several desirable relationships appeared in the spring data base, suggesting a cumulatively positive impact of SPECIALIZATION. These relationships include more reflection/innovation, more cooperation, and more interest/participation in groups with more specialized staff.

Classes with highly educated caregivers also were marked by relatively high frequencies of REFLECTION/INNOVATION on the part of children (both fall and spring), but also by low frequencies of COOPERATION (fall and spring) and less TASK PERSISTANCE (spring only). However, several of the relationships between staff education and child behavior may be due to statistical artifacts. Of all the NDCS independent variables, YEARS OF EDUCATION was most frequently affected by the biweighting procedures, and significant results were frequently eliminated or softened in the biweighted regressions for both fall and spring (6 of 14 significant results appeared to be considerably weaker in the biweighted regressions than in the ordinary regression computations). Apparently significant regression results may also be due to multicollinearily or to legitimate interactions among the predictor variables. Thus, it is difficult to draw conclusions concerning the relationship of caregiver education to child behavior in the NDCS.

The two experience variables are rarely related to child behavior. Further, no relationships persist from fall to spring or across APS and 49-center samples.

Discussion of Regression Model

Detailed discussion of the CFI findings is concentrated on data collected in spring 1977 for the entire NDCS sample of day care classrooms. Not only did spring data collection procedures minimize observer effects, but the data themselves were likely to reflect patterns of child behavior that have stabilized over the year. The discussion of regression results includes consideration of consistency between the fall and spring data analyses and between the 49 center and APS samples of classrooms. Results of the relevant standard and biweighted regressions are shown for each of the dependent variables separately for free play activities and adult-directed activities (see Tables 6.2

to 6.16) Individual coefficients for the covariables are not reported in the regression tables since there was no evidence of interaction between the covariables and the policy variables. However, significant relationships are reported in the text and the contribution of the covariables to the total R^2 is indicated in the tables.

Reflection/Innovation (Table 6.2). In both observation contexts--free play and adult-directed activity--more REFLECTION/INNOVATION on the part of children was associated with smaller groups. This relationship was particularly strong in the APS sample ($r = -.54$, $p<.01$ for free play; $r = -.50$, $p<.01$ during adult-directed period) but similar trends were also found in the 49-center sample. Also, a significant relationship indicating more REFLECTION/INNOVATION in small groups was found in the fall data base during adult-directed activities but not during free play.

More YEARS OF EDUCATION for caregivers was associated with more REFLECTION/INNOVATION for children during free play (at both fall and spring data collection points). This result is found primarily in the 49-center sample. During adult-directed periods, a similar relationship exists for specialized caregiver education/training, but only in the spring. At that time, more SPECIALIZATION is related to more REFLECTION/INNOVATION in both the APS ($r = .55$, $p<.01$) and the 49-center sample ($r = .21$, $p<.05$). Finally, more CAREGIVER EXPERIENCE is related to more child REFLECTION/INNOVATION during free play in the spring.

Verbal Initiative (Table 6.3). During the fall observations, older children offered more opinions than younger children for both free play and adult-directed activities. This result was found in both 49-center and APS samples at that time. No relationship between age of children and frequency of verbal initiative was found in the spring data collection.

At both data collection points and for both free play and adult related activities, children offered more opinions when in small groups. This trend was consistent, though not always significant, in both the APS and the 49-center groups.

YEARS OF EDUCATION was related to VERBAL INITIATIVE in the fall observations only. At that time, children in classrooms with more education expressed opinions more often. This relationship was found primarily in the APS sample ($r = .44$, $p<.05$ for free play, and $r = .52$, $p<.01$ for adult-directed activity).

Cooperation (Table 6.4). During the fall observations, COOPERATION was related to the covariables during both free play and adult-directed activities. Older children were more cooperative during adult-directed activities and children from low SES families were slightly more cooperative during free play. These relationships did not recur in the spring observations.

While no group composition measures were related to cooperation in the fall, children in smaller groups cooperated relatively more often in the spring regardless of the activity. This trend was found in both the APS and 49-center samples.

In the spring, COOPERATION during free play activities was related to caregiver SPECIALIZATION. More cooperation was associated with higher proportions of caregivers with specialized education/training. YEARS OF EDUCATION was a significant regressor for free play activities, in both the fall and spring, but, since the simple correlation of EDUCATION and COOPERATION was essentially zero, this apparent effect is probably due either to multicollinearity or to interactions among the policy variables. The fact that this result was true for the 49-center sample but not the APS also supports the nonsignificance of an education effect. None of the other caregiver qualifications was associated with COOPERATION during adult-directed activities.

Non-Involvement (Table 6.5). Higher SES groups tended to have more non-involvement in both activity contexts and at both data points. In the spring, child NON-INVOLVEMENT during free play activities was significantly related to GROUP SIZE: non-involvement tended to be more frequent in larger classrooms. In the context of adult-directed activities, child NON-INVOLVEMENT was not related to GROUP SIZE. However, there was more NON-INVOLVEMENT in lower ratio classrooms.

The level of non-invvolvement in a classroom was negatively related to caregiver education/training in a child-related area. That is, there tended to be more activity in classrooms where more caregivers had specialized preparation. This result held for all activity types in both APS and 49-center samples in the fall, but was evident for free play activities in the spring data collection.

Aimless Wandering (Table 6.6). The frequency of children aimlessly wandering in the classroom was related both to GROUP SIZE and to STAFF/CHILD RATIO in the spring observations. Wandering children were more frequently observed in larger classrooms and in classrooms with lower staff/child ratios. This pattern existed for both free play and adult-directed activities, although the GROUP SIZE effect for AIMLESS WANDERING was substantially weaker during the adult-directed activities.

Task Persistence (Table 6.7). Children from lower SES families remained in a single task longer during free play activities in the fall and during both activty types

Table 6.5

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES

Dependent Variable: Non-Involvement

Spring 1977 (N = 116)

DURING FREE PLAY ACTIVITIES

Policy Variables	Least Squares Coefficient Biweighted	Least Squares Coefficient Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
Observed group size	.003	.003	3.85	.00	.30**	.19
Observed staff/child ratio	.057	.065	0.41	.68	-.18*	(.33)
Child-related education/ training	-.022[a]	-.035	3.32	.00	-.34**	
Staff education	.008	.006	1.34	.18	.03	
Previous day care experience	-.006	-.006	1.42	.16	-.26**	
Experience in current day care center	.000	.001	0.68	.50	-.13	

DURING ADULT-DIRECTED ACTIVITIES

Policy Variables	Least Squares Coefficient Biweighted	Least Squares Coefficient Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
Observed group size	.001	.001	0.87	.39	.17	.07
Observed staff/child ratio	-.073[a]	-.150	1.59	.11	-.26**	(.15)
Child-related education/ training	-.019	-.012	1.18	.24	-.21*	
Staff education	.001	-.001	0.42	.67	-.05	
Previous day care experience	.006	.002	0.66	.51	-.03	
Experience in current day care center	.001	.000	0.10	.99	.01	

*$p<.05$

**$p<.01$

[a] indicates changes greater than 1/2 standard error of least squares coefficient in biweighted regression.

Table 6.6

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES

Dependent Variable: Wandering
Spring 1977 (N = 116)

DURING FREE PLAY ACTIVITIES

	Least Squares Coefficient					
Policy Variables	Biweighted	Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
Observed group size	.002	.002	2.17	.03	.33**	.17
Observed staff/child ratio	-.231	-.229	1.82	.07	-.30**	(.17)
Child-related education/ training	-.004	-.003	0.25	.80	-.14	
Staff education	-.004	-.004	1.01	.32	-.16	
Previous day care experience	-.006	-.006	1.21	.23	-.20*	
Experience in current day care center	-.001	-.001	0.34	.73	-.06	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	.002	.002	1.51	.13	.21*	.16
Observed staff/child ratio	-.274	-.294	2.77	.01	-.31**	(.17)
Child-related education/ training	-.003	-.007	0.18	.86	-.14	
Staff education	-.008	-.006	1.45	.15	-.16	
Previous day care experience	-.006	-.006	1.31	.20	-.20*	
Experience in current day care center	-.001	-.001	0.42	.68	-.06	

*p<.05
**p<.01

Table 6.7

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES
Dependent Variable: Task Persistence
Spring 1977 (N = 116)

DURING FREE PLAY ACTIVITIES

Policy Variables	Least Squares Coefficient Biweighted	Least Squares Coefficient Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
Observed group size	-.004	-.005	0.34	.73	-.06	.13
Observed staff/child ratio	.285	.323	1.94	.06	.25**	(.20)
Child-related education/ training	.054	.058	2.82	.01	.31**	
Staff education	-.012	-.016	2.69	.01	-.15	
Previous day care experience	.015	.014	1.97	.05	.25**	
Experience in current day care center	.001	.001	0.15	.88	.17	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	.003	.005	1.62	.11	.10	.08
Observed staff/child ratio	.311	.298	2.09	.04	.19*	(.13)
Child-related education/ training	.091	.093	2.22	.03	.21*	
Staff education	-.010	-.012	1.86	.07	-.09	
Previous day care experience	.002	.002	0.12	.91	.06	
Experience in current day care center	.002	.002	0.22	.83	.09	

*p<.05
**p<.01

in the spring. Also, for adult-directed activities in the fall and for both activity types in the spring, children remained involved in tasks longer when staff/child ratio was higher. A trend toward longer activities in larger groups was found for adult-directed activities.

The strongest relationship to TASK PERSISTENCE was found with SPECIALIZATION. Children remained in activities longer where more staff had specialized preparation This relationship was evident both during free play and adult-directed activities and at the two data collection points. For free play activities in the spring, groups where staff had more experience in other day care centers exhibited more task persistence. YEARS OF EDUCATION was a significant regressor in the spring (for both free play and adult-directed activities) but was not strongly correlated with TASK PERSISTENCE; thus, as with COOPERATION, this apparent relationship is probably due to multicollinearity or to legitimate interactions.

Interest/Participation (Table 6.8). Few relationships between the policy variables and the INTEREST/PARTICIPATION construct were discovered in the Phase III data base. In the fall, larger groups were associated with the lower participation. The only relationship found in the spring, however, was a weak tendency for more specialized staff to have groups with a greater level of interest/participation.

Classroom Activity Balance (Table 6.9). The dominant relationships found for the structural balance of activities in the classroom were for the covariables. Younger children and children from higher SES families participated more often in open, unstructured activities both in fall and spring. In addition, the incidence of

Table 6.2

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES

Dependent Variable: Reflection/Innovation
Spring 1977 (N = 116)

	Least Squares Coefficient			Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
Policy Variables	Biweighted	Ordinary	t			
			DURING FREE PLAY ACTIVITIES			
Observed group size	-.0029	-.004	1.75	.08	-.25**	.13
Observed staff/child ratio	.007	.013	0.14	.90	.11	(.13)
Child-related education/ training	-.003	-.009	0.89	.38	.03	
Staff education	.001[a]	.005	1.62	.11	.13	
Previous day care experience	.003	.004	1.28	.20	.12	
Experience in current day care center	.004	.004	2.44	.02	.20*	
			DURING ADULT-DIRECTED ACTIVITIES			
Observed group size	-.001	-.001	1.71	.09	-.19*	.08
Observed staff/child ratio	.045	.014	0.12	.90	.12	(.11)
Child-related education/ training	.014	.018	1.91	.06	.28**	
Staff education	.000	.002	0.49	.63	.12	
Previous day care experience	.002	.001	0.30	.76	.10	
Experience in current day care center	.001	.001	0.56	.58	.14	

*$p < .05$
**$p < .01$

[a] indicates change greater than 1/2 standard error of least squares coefficient in biweighted regression.

Table 6.3

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES
Dependent Variable: Verbal Initiative
Spring 1977 (N = 116)

DURING FREE PLAY ACTIVITIES

Policy Variables	Least Squares Coefficient Biweighted	Least Squares Coefficient Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
Observed group size	-.001	-.001	2.12	.04	-.21*	.11
Observed staff/child ratio	.049	.109	1.40	.17	.15	(.16)
Child-related education/ training	.003	.006	0.62	.54	.24*	
Staff education	.003	.004	1.43	.16	.24*	
Previous day care experience	-.004	-.003	1.01	.32	.01	
Experience in current day care center	-.002	-.002	1.40	.16	-.10	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	-.001	-.001	1.87	.06	-.19*	.08
Observed staff/child ratio	.067	.047	0.55	.58	.08	(.11)
Child-related education/ training	.003	.003	0.30	.76	-.01	
Staff education	.004	.003	1.06	.29	.19*	
Previous day care experience	-.002	-.004	1.22	.22	-.04	
Experience in current day care center	-.001	-.002	1.40	.16	-.15	

*$p < .05$
**$p < .01$

Table 6.4

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES
Dependent Variable: Cooperation
Spring 1977 (N = 116)

Policy Variables	Least Squares Coefficient Biweighted	Least Squares Coefficient Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
DURING FREE PLAY ACTIVITIES						
Observed group size	-.007[a]	-.006	2.44	.01	-.24**	.11
Observed staff/child ratio	-.038	-.097	0.26	.79	.08	(.13)
Child-related education/ training	.112	.137	2.97	.00	.22*	
Staff education	-.013[a]	-.026	1.97	.05	-.08	
Previous day care experience	.001	-.004	0.24	.81	.07	
Experience in current day care center	-.010	-.009	1.17	.24	.04	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	-.005	-.005	1.87	.06	-.21*	.06
Observed staff/child ratio	.141	.147	0.46	.65	.13	(.07)
Child-related education/ training	.036	.042	1.18	.24	.11	
Staff education	-.012	-.003	0.30	.76	.07	
Previous day care experience	-.006	.003	0.28	.78	.10	
Experience in current day care center	-.007	-.003	0.47	.64	-.02	

*$p<.05$
**$p<.01$

[a] indicates changes greater than 1/2 standard error of least squares coefficient in biweighted regression.

Table 6.8

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES

Dependent Variable: Interest/Participation

Spring 1977 (N = 116)

DURING FREE PLAY ACTIVITIES

Policy Variables	Least Squares Coefficient		t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
	Biweighted	Ordinary				
Observed group size	-.002	-.003	1.20	.28	-.15	.06
Observed staff/child ratio	.242	.285	0.92	.35	.17	(.08)
Child-related education/training	.054	.053	1.28	.12	-.14	
Staff education	-.003	-.004	.25	.70	-.15	
Previous day care experience	.002	.002	.02	.90	-.19	
Experience in current day care center	-.004	-.004	.78	.47	-.07	

$^*p<.05$
$^{**}p<.01$

Table 6.9

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES

Dependent Variable: Classroom Activity Balance
Spring, 1977 (N = 116)

ALL ACTIVITIES COMBINED

Policy Variables	Ordinary Least Squares Coefficient		t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
Observed group size	.003	.003	1.36	.21	.13	.03
Observed staff/child ratio	.100	.140	0.35	.68	.08	(.21)
Child-related education/ training	-.035	-.031	0.83	.41	-.21	
Staff education	-.020	-.026	0.29	.82	-.03	
Previous day care experience	-.004	-.004	0.31	.78	-.10	
Experience in current day care center	.005	.006	0.93	.35	-.07	

*$p < .05$
**$p < .01$

unstructured activity was negatively related to STAFF/CHILD RATIO and SPECIALIZATION and positively related to YEARS OF EXPERIENCE and CENTER EXPERIENCE.

Patterns of Interaction (Tables 6.10-6.13). Orientation to the social and physical environment was associated broadly with the covariables during both fall and spring data collection points. Younger children were more frequently oriented to adults and less often to individual children or groups regardless of activity type. In the fall, children from higher SES families were more often oriented towards the physical environment than were those from lower SES families.

Both in the fall and spring, children in small groups attended to adults more frequently and towards groups of children less frequently. The relationships between caregiver qualifications and the child's patterns of interactions were extremely scattered.

Transitional Activities (Tables 6.14-6.16). The codes associated with various activity transitions (MONITORS ENVIRONMENT, MOVES WITH PURPOSE and GROUP PASSIVE ACTIVITY) were associated with the covariables in the fall but not the spring. In the fall, younger children moved purposefully less often than they engaged in a greater degree of environmental monitoring (each in free play periods only).

Relationships of these variables with the size of the group appeared in both fall and spring data. Children in large groups were more often engaged in environmental monitoring regardless of activity type, spent more time in adult-directed passive activities (spring only) and more often moved with purpose (fall only). As with patterns of interaction, this group of codes was rarely and inconclusively related to caregiver qualifications.

Table 6.10

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES

Dependent Variable: Orientation to Adults

Spring 1977 (N = 116)

Policy Variables	Least Squares Coefficient		t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
	Biweighted	Ordinary				
DURING FREE PLAY ACTIVITIES						
Observed group size	-.008	-.006	3.11	.00	-.30**	.11
Observed staff/child ratio	.159	.231	0.84	.40	.18*	(.15)
Child-related education/ training	.015	.028	0.80	.43	-.01	
Staff education	-.011	-.009	0.94	.35	-.05	
Previous day care experience	-.002	-.004	0.31	.76	.07	
Experience in current day care center	-.007	-.009	1.41	.16	-.18*	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	-.007	-.007	2.78	.01	-.27**	.11
Observed staff/child ratio	-.120	-.105	0.32	.75	.08	(.14)
Child-related education/ training	.026	.026	0.74	.46	.02	
Staff education	-.021	-.000	0.02	.96	.09	
Previous day care experience	.008	.009	0.78	.44	.13	
Experience in current day care center	-.002	-.002	0.34	.74	-.07	

*p<.05
**p<.01

Table 6.11

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES
Dependent Variable: Orientation to Children
Spring 1977 (N = 116)

DURING FREE PLAY ACTIVITIES

Policy Variables	Least Squares Coefficient Biweighted	Least Squares Coefficient Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
Observed group size	.002	.002	1.17	.25	.14	.04
Observed staff/child ratio	-.094	-.078	0.41	.68	-.09	(.06)
Child-related education/ training	-.003	-.004	0.16	.87	.05	
Staff education	.006	.006	0.96	.34	.09	
Previous day care experience	-.001	-.001	0.08	.93	-.03	
Experience in current day care center	.002	.002	0.54	.59	.04	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	.003	.003	1.33	.19	.10	.08
Observed staff/child ratio	.101	.121	0.47	.64	-.01	(.18)
Child-related education/ training	-.002	-.002	0.07	.94	.15	
Staff education	.013	.016	1.99	.05	.28**	
Previous day care experience	-.000	-.000	0.02	.99	.02	
Experience in current day care center	.001	.001	0.17	.86	.03	

*$p < .05$
**$p < .01$

Table 6.12

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES
Dependent Variable: Orientation to Groups
Spring, 1977 (n = 116)

DURING FREE PLAY ACTIVITIES

Policy Variables	Least Squares Coefficient		t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
	Biweighted	Ordinary				
Observed group size	.003	.003	2.86	.00	.28**	.11
Observed staff/child ratio	.185	.172	1.58	.12	.14	(.16)
Child-related education/ training	.008	.008	0.54	.59	.05	
Staff education	-.004	-.003	0.90	.37	-.12	
Previous day care experience	.000	.000	0.04	.97	-.01	
Experience in current day care center	-.005	-.004	1.40	.16	-.09	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	.005	.004	2.43	.02	.29**	.12
Observed staff/child ratio	-.010	-.017	0.20	.84	-.04	(.14)
Child-related education/ training	.012	.012	0.68	.53	.01	
Staff education	-.005	-.006	1.25	.21	-.14	
Previous day care experience	.003	-.004	0.54	.59	-.07	
Experience in current day care center	-.009	-.008	2.07	.03	-.16	

*$p<.05$
**$p<.01$

Table 6.13

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES
Dependent Variable: Orientation to Environment
Spring 1977 (N = 116)

Policy Variables	Least Squares Coefficient Biweighted	Least Squares Coefficient Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
DURING FREE PLAY ACTIVITIES						
Observed group size	-.003	-.003	1.72	.22	-.20*	.06
Observed staff/child ratio	.040	.036	0.79	.44	-.00	(.06)
Child-related education/ training	.003	.002	0.49	.62	.06	
Staff education	.004	.003	0.70	.48	.10	
Previous day care experience	-.022	-.-18	1.60	.11	-.16	
Experience in current day care center	.000	.000	0.06	.94	.02	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	-.002	-.002	0.83	.41	-.05	.04
Observed staff/child ratio	-.050	-.045	0.15	.89	-.01	(.04)
Child-related education/ training	.010	.002	0.07	.95	0.04	
Staff education	.002	-.003	0.32	.76	-.05	
Previous day care experience	-.010[a]	-.022	1.64	.10	-.16	
Experience in current day care center	.001	-.000	0.06	.96	.02	

*p<.05
**p<.01

[a] indicates changes of greater than 1/2 standard error of least squares coefficient in biweighted regression.

Table 6.14

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES
Dependent Variable: Monitors Environment
Spring 1977 (N = 116)

Policy Variables	Least Squares Coefficient Biweighted	Least Squares Coefficient Ordinary	t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
DURING FREE PLAY ACTIVITIES						
Observed group size	.001	.001	1.51	.14	.14	.04
Observed staff/child ratio	.018	.022	0.25	.82	-.06	(.06)
Child-related education/ training	-.012	-.011	0.83	.41	-.13	
Staff education	-.002	-.002	0.57	.58	-.09	
Previous day care experience	-.001	-.001	0.18	.86	-.08	
Experience in current day care center	.001	.001	0.25	.73	-.03*	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	.002	.002	1.87	.06	.24*	.09
Observed staff/child ratio	-.090	-.117	0.88	.38	-.15	(.09)
Child-related education/ training	.015	.019	1.34	.19	.06	
Staff education	-.004	-.005	1.29	.21	-.10	
Previous day care experience	.004	.004	0.87	.39	.04	
Experience in current day care center	-.000	-.001	0.15	.83	-.03	

*$p<.05$
**$p<.01$

Table 6.15

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES

Dependent Variable: Moves with Purpose
Spring 1977 (N = 116)

DURING FREE PLAY ACTIVITIES

Policy Variables	Least Squares Coefficient		t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
	Biweighted	Ordinary				
Observed group size	.001	.000	1.39	.17	.14	.07
Observed staff/child ratio	.030	.078	0.54	.50	.05	(.09)
Child-related education/ training	.006	.008	0.59	.48	.18	
Staff education	.002	.003	1.45	.14	.16	
Previous day care experience	-.002	-.002	-.93	.36	-.05	
Experience in current day care center	-.001	-.001	0.51	.46	-.02	
DURING ADULT-DIRECTED ACTIVITIES						
Observed group size	.000	.000	0.45	.65	.02	.07
Observed staff/child ratio	-.029	.056	1.14	.26	.12	(.11)
Child-related education/ training	-.002	.004	0.60	.47	.11	
Staff education	.002[a]	.003	1.75	.08	.22*	
Previous day care experience	.001	.001	0.16	.88	.06	
Experience in current day care center	-.000	-.000	0.54	.64	-.09	

*$p < .05$
**$p < .01$

[a] indicates changes of greater than 1/2 standard error of least squares coefficient in biweighted regression.

Table 6.16

RESULTS OF REGRESSIONS OF CHILD BEHAVIOR VARIABLES ON SELECTED POLICY VARIABLES

Dependent Variable: Group Passive Activity

Spring 1977 (N = 116)

DURING FREE PLAY ACTIVITIES

Policy Variables	Least Squares Coefficient		t	Significance of t	Simple Correlation	R^2 for Policy Variables (R^2 with Covariables)
	Biweighted	Ordinary				
Observed group size	.002	.002	1.37	.15	.25**	.15
Observed staff/child ratio	-.029	-.032	0.66	.51	-.06	(.18)
Child-related education/ training	-.020	-.024	1.28	.20	-.15	
Staff education	.006	.005	1.36	-.18	-.19*	
Previous day care experience	.003	.003	0.83	.42	.08	
Experience in current day care center	-.002	-.002	2.08	.04	-.30**	

*$p<.05$
**$p<.01$

Determinants of Rare but Important Events

Some of the CFI codes that occurred infrequently (e.g., only a few times per thousand frames of observation) might be viewed as having unusual psychological importance or as being unusually revealing regarding the behavioral climate of a day care center. Relevant codes, termed "critical incidents" are listed in Table 6.17, along with their frequencies of occurrence in Fall 1976 and Spring 1977. Because of their rarity and because a code that was recorded once tended to recur over several frames, these events exhibited extremely skewed distributions across classrooms, with many classes showing no occurrences of a given behavior and other classes showing small flurries of critical events, e.g., a brief hostile exchange between children, followed by a few minutes of crying.

Ordinary regression analysis embodies distributional assumptions that are violated by rare events of this kind. However, logit analysis, an alternative form of regression, is designed to handle such events. In essence, logit analysis estimates the odds of a rare event occurring at all in a given classroom, characterized by a given configuration of policy variables. (In contrast, ordinary regression as it has been used elsewhere predicts the frequency of a given event as a function of policy variables.)

A series of logit analyses was conducted, using as dependent variables the eighteen rare codes listed in Table 6.17. The independent variables, which were similar to those used in the regression analysis, included STAFF/CHILD RATIO, GROUP SIZE, YEARS OF EDUCATION, PREVIOUS EXPERIENCE, CENTER EXPERIENCE, and two covariables--CHILD AGE and STAFF AGE. Analyses were conducted separately for fall and spring, as well as for the Altanta Public Schools and each of the 49-center sites. Thus, for each pairing of

Table 6.17

FREQUENCIES OF CRITICAL INCIDENTS CODES AS PERCENTAGE OF ALL CODES

	Fall 1976	Spring 1977
Offers sympathy	0.1	0.0
Shares, helps	0.6	0.6
Receives praise	0.4	0.3
Asks for comfort	0.1	0.0
Receives comfort	0.3	0.3
Crying	0.2	0.2
Avoids, withdraws	0.1	0.1
Isolates self	0.7	0.1
Hostile exchange	0.1	0.1
Intrudes hostilely	0.2	0.1
Receives hostile intrusion	0.1	0.1
Receives rejection	0.1	0.1
Refuses to comply	0.3	0.2
Hostilely asserts rights	0.1	0.1
Temper tantrum	0.0	0.1
Receives threats	0.4	0.3
Receives physical punishment	0.0	0.0
Experiences accident	0.1	0.0

an independent variable with a dependent variable (108 such pairs in all) there were eight separate opportunities for a positive or negative relationship to appear (four sets of centers at two different time points).

Given this situation, it is not surprising that the pattern of outcomes is exceedingly complex. Relatively few relationships achieve conventional levels of statistical significance taken in isolation. However, the primary concern was not with relationships occurring in a particular place at a particular time but with broader relationships that were fundamentally invariant across places and times. To identify such relationships the following criteria for declaring the existence of "consistent" effects were adopted:

(1) The signs of coefficients were consistently positive (or negative) in all, or in all but one, of the possible cases, and

(2) Either the inconsistent coefficient was not significant at the .05 level

(3) Or at least one of the consistent coefficients was significant at the .05 level.

Table 6.18 summarizes the results of applying these criteria to the array of data generated by the multiple logit analyses. The table is in the form of a matrix of dependent variables (rare codes) crossed by independent variables (policy variables and covariables). Wherever a "+" sign appears in a cell at the intersection of a particular dependent or independent variable, it indicates that a consistent positive association was found using the definition above. A "-" sign, analogously, indicates a consistent negative relationship. An asterisk in a cell indicates that at least one coefficient was significant at the .05 level. (For technical reasons, logit analyses were not possible in all eight cases for every variable. In some cases, events occurred at least once in every classroom or in a particular

Table 6.18

Relationships Between Policy Variables and Critical Incidents[a]

	Child Age	S/C Ratio	Group Size	Staff Educ.	Staff Exper	Staff Age	N
Offers sympathy							8
Shares, helps							0
Receives praise	+	+	+	+	-	+	3
Asks for comfort	-						8
Receives comfort		+		-			4
Crying				-			8
Avoids, withdraws		+*		+			7
Attends self			+				7
Hostile exchange			+*				8
Intrudes hostilely	+		+	-			6
Receives hostile intrusion			+	-*			8
Receives rejection	+*			-			8
Refuses to comply		+		+			6
Hostilely asserts rights							7
Temper tantrum							7
Receives threats		+	+				4
Receives physical punishment			+*				6
Experiences accidents							8

[a] Cell entries--"+" of "-" signs--indicate directions of consistent relationships.

*Indicates significance at the .05 level in at least one case.

sample. Numbers listed in the right-hand column (N) of Table 6.18 indicate the number of analyses on which each consistency judgment is based.

Though this method of assessing consistency is approximate at best, the results are revealing. Large groups are associated with indices of conflict (HOSTILE EXCHANGE, INTRUDES HOSTILELY, RECEIVES THREATS, RECEIVES PHYSICAL PUNISHMENT, RECEIVES HOSTILE INTRUSION) and of withdrawal (ATTENDS SELF). In only one case (RECEIVES PRAISE) are large groups associated with a critical event that would generally be regarded as positive. High STAFF/CHILD RATIOs are associated with two categories of experience that might be regarded as beneficial to children (RECEIVES COMFORT, RECEIVES PRAISE), but also with other categories that might be seen as negative (RECEIVES THREATS, AVOIDS, WITHDRAWS and REFUSES TO COMPLY). High levels of STAFF EDUCATION are associated with low likelihood of CONFLICT and REJECTION and high likelihood of PRAISE, but also high likelihood of AVOIDANCE/WITHDRAWAL and REFUSAL TO COMPLY. Once again, small groups are associated with a pattern of outcomes that, in our view, is more consistently desirable than the patterns associated with any other policy variables. In contrast, high staff/child ratios seem to be associated with a general intensification of emotional relationships, i.e., with relatively extreme expressions of both warmth and anger. The highly educated caregiver appears to have a distinctive style, marked by avoidance of conflict. However, because the critical incident analysis was pursued independently of other portions of the analyses, no attempt was made to separate effects of education from those of specialization in a child-related field.

Child Behavior in Structured Situations

It has been shown that some behaviors of psychological interest occur infrequently in natural settings because of a simple lack of opportunity for children to act in ways that meet the definitions of relevant observation codes. Historically this has been one major reason why so much developmental research takes place in contrived laboratory settings. The legitimate intent of this kind of research has been to achieve maximum control over relevant variables, i.e., standardization of situations to which all subjects are exposed, and exclusion of extraneous influences of various kinds. To achieve such control, ecological validity has often been sacrificed.

The behavior of children and adults in natural settings reflects both the enduring characteristics of the individuals involved and the opportunities and incentives provided by the situation. Thus, for example, when an angry child is observed, it is difficult to determine without additional information whether the child is characteristically belligerent or whether his/her anger is due to some frustration, threat or infringement of rights. For NDCS purposes, measurement of situational pressures was more important than measurement of child traits. Nevertheless, the character traits of children might change as a consequence of different types of care, and this trait measurement was of interest as well (even though only short-term change could potentially be studied during the NDCS).

The kinds of situations in which children find themselves--stimulating or boring, threatening or reassuring--are part of the day care process and may also be affected by such classroom characteristics as staff/child ratio, group size and staff qualifications. Since those characteristics of classrooms vary systematically in the NDCS, the opportunity for observation of behavior associated with child traits may vary as well. During Phase II of the NDCS, several important child bahaviors occurred infrequently in the course of naturalistic observation. Consequently, an effort was made to enhance Phase III observation data by using contrived play situations designed to highlight selected child behaviors toward peers and play materials.

In both Fall 1976 and Spring 1977 same-sex, same-age pairs of children were placed in two structured situations, intended to present clear opportunities for certain types of behavior that were relatively rare in natural settings and that--if influenced by the policy variables--would represent important domains of effects. The situations provide the opportunity, but not necessity, for voluntary cooperation and sharing, and for creative and cooperative use of materials. The two structured situations were arranged as follows:

- In the limited resources situation, the children were given a Play-Doh game with one Play-Doh mold but an abundant quantity of Play-Doh. The crux of the situation was that only one child could use the mold.
- In the abundant resources situation, the children were given a Fisher-Price Play Family Village and associated materials. This toy permits independent play, cooperative play, and mutual fantasy play.

In both cases, behavior was recorded using the standard CFI. The structured situations achieved their goal of altering the frequencies of certain important forms of behavior (see Table 6.19). For example, frequencies of open-ended, cooperative play, innovative use of materials and reflective behavior all increased dramatically.

Upon examination, the patterns of intercorrelations for frequently used codes were found to be similar for the two structured observation segments during both data collection periods. Correlations for codes from one segment to the other, and correlations for codes from one data collection period to the other, are shown in Table 6.20. The pattern across segments is one of strong relationship with three exceptions: CONSIDERS, CONTEMPLATES PROBLEMS, SEES PATTERN,

Table 6.19

FREQUENCIES OF STRUCTURED OBSERVATION CODES[a]
(Fall, 1976 and Spring, 1977 Combined)

	Classroom Observations	Structured Observation	
		Limited Resources	Abundant Resources
A. Activity Codes			
Shares open, expressive activity	12.8	26.0	30.1
Monitors environment	10.5	24.3	5.8
Offers opinions	7.8	7.6	10.5
Receives general comments (from other child)	2.8	4.2	5.1
Receives demands, requests (from other child)	0.6	2.8	1.7
Receives information (from other child)	0.2	1.9	1.3
Adds prop or idea	2.7	10.7	27.1
Considers, contemplates problem	1.6	2.1	7.9
Gives orders, directs others (child)	1.0	2.2	1.9
Receives request to share (from other child)	0.1	2.9	0.7
Sees pattern, solves problem	0.1	0.4	1.2
B. Object Codes			
Attention to child	23.0	49.1	31.7
Attention to environment	41.9	44.7	66.4
C. Activity Continuity Codes			
Not involved in activity	7.3	21.9	0.2

---

[a] Tabled codes are those which exceed one percent frequency for either limited or abundant resource segments. Code frequencies are shown as a percentage of all observation codes. For both behavior and object codes, the total number recorded was approximately 725,000 (classroom observations), 40,000 (limited resources), and 70,000 (abundant resources).

Table 6.20

STABILITY OF STRUCTURED OBSERVATION CODES

	CORRELATION BETWEEN LIMITED RESOURCES AND ABUNDANT RESOURCES		CORRELATION BETWEEN FALL, 1976 AND SPRING, 1977 DATA COLLECTION PERIODS	
			SEGMENTS	
CODE DESCRIPTION	LIMITED RESOURCES	ABUNDANT RESOURCES	Fall 1976	Spring 1977
Open, unstructured activity	.14	.13	.38	.64
Monitors environment	.39		.43	.51
Verbal initiative	.34	.39	.75	.55
Receives general comments	.12		.36	.39
Receives demands, requests	.12		.36	.39
Receives information			.43	.68
Adds prop or idea			.32	.55
Considers, contemplates problem			.11	.11
Sees pattern, solves problem			.18	.15
Gives orders, directs others		.22	.51	.61
Receives request to share			.14	.31
Attention to Child	.17	.33	.33	.60
Attention to Environment	.19	.33	.37	.60
Not involved in activity	.13			

Note: all correlations shown are $p < .15$; T3: n = 117; T4: n = 116; T3 to T4: n = 111.

SOLVES PROBLEM, NOT INVOLVED IN ACTIVITY. Significant stability across data collection periods was also evident.

Eight structured observation measures (all but one combined across limited and abundant resource segments) were entered into regression analyses with a set of predictors identical to that used in the classroom observation analyses. Two covariables (CHILD AGE, FAMILY SES) and six independent variables (GROUP SIZE, STAFF/CHILD RATIO, SPECIALIZATION, YEARS OF EDUCATION, PREVIOUS DAY CARE EXPERIENCE and EXPERIENCE) were entered, and the APS and 49-center samples were analyzed separately. The second covariable (SES) was not used for the APS sample since since that sample was economically and racially homogeneous.

The regression results are summarized in Table 6.21. Two points relating to these analyses are worthy of comment. First, the covariables were somewhat stronger predictors of child behavior in the structured observations than were the NDCS policy variables. The relative strength of CHILD AGE as a predictor was weaker in the spring than in the fall, as was the case in the classroom observations. FAMILY SES is a relatively stronger predictor in the spring data. No consistant or clearly powerful relationships were observed between structured situation behavior and the NDCS policy variables.

Second, replication of effects were extremely rare. For both fall and spring, older children exhibited more verbal initiative and engaged in less environmental monitoring than younger children, and children from higher SES families offered more opinions than those from lower SES families. Only two relationships were replicated across APS and 49-center samples: in the fall, children from groups where staff had more experience in other day care centers made more requests; in the spring, children from small groups spent more time considering materials/solving problems.

Table 6.21

Summary of Significant Structured Situation Regression Results

	Open, Unstructured Activity (Individual Play)	Monitors Environment	Verbal Initiative	Child Requests	Adds Prop, Idea	Reflection/ Innovation (limited)	Reflection/ Innovation (amundant)	Orientation to Child
Fall Observation Period								
Covariables								
Age of Children		-	+	(+)				+
SES of Family	-		+					+
Group Composition								
Group Size								
Staff/Child Ratio								
Caregiver Qualifications								
Specialization						(-)		
Years of Education		(+)		-				
Previous Day Care Experience				+[a]				
Experience in Current Center								
Spring Observation Period								
Covariables								
Age of Children		-	(+)					
SES of Family	+	-	(+)		-	(-)	-	
Group Composition								
Group Size					(-)	(-)[a]		+
Staff/Child Ratio				(+)				
Caregiver Qualifications								
Specialization				(+)	-			
Years of Education								
Previous Day Care Experience								
Experience in Current Center					+			

Results noted were significant at $p<.05$; results significant at $.05<p<.15$ shown in parentheses.

[a] results significant at $p<.15$ in both APS and 49-center samples.

For NDCS purposes the important conclusion to be drawn from this set of results is that effects of the policy variables are very much tied to the classroom situation. In more-or-less standardized situations the CFI measures tend to reflect powerful and enduring influences of general developmental status and family background. When used in natural group settings, the CFI captures group dynamics that are subject to influence by certain regulatable center characteristics.

CHAPTER SEVEN: CHILD OUTCOMES AND CHILD CARE POLICY

Unlike most previous studies of the impact of group care on young children, the National Day Care Study (NDCS) had a broad mandate: to trace the connections between a wide range of regulatable center characteristics on the one hand and both quality and costs on the other. Of the three domains, *quality* clearly poses the most difficult conceptual and technical problems. Philosophical disagreements as to its definition abound; instruments adequate to measure all of its aspects do not exist.

Effects and *quality* are not identical but are intimately linked. The effects of different day care center arrangements are, in principle, objective matters, although they may be difficult to conceptualize and measure. Assessing quality requires that value judgments be placed on different patterns of effects. There are many open questions in the developmental research literature and in the practical lore of child care regarding the objective effects of different center characteristics.

A major assumption of the NDCS was that *judgments of quality should rest on direct assessment of the child's daily experience* and its developmental impact. Such judgments could not be based simply on characteristics of centers such as numbers of staff and children, background qualifications of staff, or features of the physical environment. Indeed, the whole purpose of the NDCS effects analyses was to determine whether such center characteristics have the kinds of impacts on day care processes and outcomes assumed by current federal purchasing standards, the FIDCR, and some state and local licensing requirements.

CFI findings show that certain regulatable center characteristics are consistently associated with patterns of child behavior and, by inference, with the quality of day care settings. Smaller groups of children and caregivers and, to a lesser extent, higher staff/child ratios are associated with more desirable classroom behavior. Moreover, aspects of caregiver qualifications, especially education/training relevant to young children, are associated with positive classroom dynamics.

- In smaller groups, as contrasted to larger ones,
  - --children show more cooperation, verbal initiative and reflective/innovative behavior; and
  - --children show less hostility and conflict and are less frequently observed to wander aimlessly or to be uninvolved in tasks or activities.
- In higher-ratio classes, as opposed to those with lower staff/child ratios,
  - --children show greater task persistence and are less frequently observed to wander or be uninvolved in activities.
- In classes supervised by lead teachers with child-related education/training
  - --children show more reflection/innovation, cooperation, greater task persistence and interest/participation and are less frequently uninvolved in tasks or activities.

The NDCS included a variety of measures and data collection instruments (Bache, 1980). SRI selected two observation systems--one focused on children and one on caregivers--to record daily behavior in considerable detail. In addition, children were tested at two points in the year using two tests of school-performance related skills, the Pre-School Inventory (PSI) and the Peabody Picture Vocabulary Test (PPVT). For these tests, the gain in child performance over the year was the measure analyzed (Goodrich and Singer, 1980).

Analyses were conducted to explore links between child behavior and changes in PSI and PPVT scores, and to gain additional insight into the effects of regulatable characteristics on behavior in the classroom. These analyses showed that certain caregiver and child behaviors, which are influenced by group size and/or child-related education/ training of staff, in turn are associated with particular rapid gains on one or both of the two tests, even with the policy variables controlled (Singer, 1980).

In centers where children had high levels of cooperation, PSI gains were large ($r=.42, p<.01$). Similarly, PSI gains were large in centers where children were frequently engaged in reflective, innovative behavior ($r=.43, p<.01$). In centers were children were allowed to wander aimlessly, uninvolved in tasks and activities, PSI gains were small ($r=.32, p<.05$). PSI gains were also somewhat sensitive to the degree of structure in classroom activities. Centers where structured activities predominate showed higher gains than centers where unstructured activities predominate ($r=.24, p<.10$).

All of the above relationships are observed even when group size is held constant. Together with group size, various combinations of the child behaviors just listed explain up to 40 percent of total center-to-center variation in PSI gains (Singer, 1980). In light of the reliabilities of the various measures involved, this fraction presents a very substantial proportion of the variation (Bache, 1980a). These findings help put NDCS findings on test performance into perspective, showing that rapid gains do not somehow occur automatically in small classes. Rather, rapid gains can be expected where children are active and engaged. Activity and engagement are more common in small groups than in large ones--and when they occur, the prognosis for test performance is good.

Gains on the PPVT are linked to somewhat diferent aspects of classroom dynamics than are PSI gains. As with PSI, reflective, innovative behavior of children was associated with rapid gains on the PPVT (r=.32,p<.05), and aimless wandering was associated with slower gains (r=.32, p<.05). However, several caregiver behaviors that showed no relationship to PSI gains did appear to affect PPVT gains: Gains were more rapid in centers where caregivers showed high levels of interaction with children, both social (questioning, instructing, responding, praising and comforting) and managerial (commanding and correcting). Gains were slower in centers where caregivers spent relatively high proportions of their time observing children. When various caregiver behaviors were combined with information on the size of subgroups around the caregiver, regression equations accounted for almost half of the center-to-center variation in PPVT gains (Singer, 1980). Comparison of the results for the PSI and PPVT suggests that PSI gains may have depended more on the child's own activities, while PPVT gains were more responsive to interaction with the caregiver.

It is natural to suppose that policy conclusions should be used primarily on very strong statistical relationships and that weak relationships constitute a less useful basis for policy decisions. However, strengths of association can be misleading, and weak relationships can be informative for policy, partly because technical factors conspire against statistically strong findings in large scale studies such as the NDCS, and partly because statistical strength does not necessarily reflect substantive importance to children.

Previous national evaluations of outcomes associated with different levels of resource outlay in education or with program variations within early intervention projects have typically shown that measured effects of variations in programs or resources are small. Often effects due to program differences cannot be detected at all. This may be true for all of several reasons. Outcome measures or research designs may lack the sensitivity to detect effects of variations in programs or resources. Program differences may be poorly defined, implemented or measured. Finally, differences in outcomes may be genuinely small relative to the effects of the program per se (e.g., of school itself, or of early intervention itself) or relative to the effects of other factors such as the socioeconomic backgrounds of the children served.

The NDCS had to contend with some of the limits of instrumentation and design that confronted earlier studies. Some measures, in particular the CFI, were less reliable than would have been ideal, and some aspects of children's development could not be measured. For statistical reasons, limited reliability of measures leads to systematic underestimation of the true strengths of relationships and can prevent genuine relationships from being detected. Moreover, because many center characteristics had to be studied as they occurred naturally, these characteristics were not fully independent of each other, creating some difficulty in disentangling their effects. The NDCS was, in effect, a search for signals in a noisy environment, as is often the case in large-scale policy studies.

On the other hand, the NDCS had advantages over previous studies. It's independent variables, particularly those relating to classroom composition (number of children and number of caregivers), were defined clearly and measured precisely. In addition, the study's dependent variables included observational measures of an unusually broad range of child and caregiver behaviors, as well as test scores. Therefore, the opportunity to detect effects in this study was significantly greater than in studies that use narrow sets of outcome measures or that focus on effects of programs that may have differed less in reality than in theory.

In sum, the NDCS had advantages over previous studies but, nevertheless, faced important technical constraints. It was not at all certain that the study would be able to find systematic differences among classrooms or centers, or that such differences, if found, would be linked to any of the study's policy variables. For all of these reasons, it can be argued that the magnitudes of the various relationships obtained are secondary. The most important aspects of the NDCS effects' findings are the following: (1) the *existence* of statistically significant relationships between policy variables and day care quality, reflected in caregiver behavior, child behavior, and test scores; (2) the *relative strengths* of different policy variables--the fact that some policy variables are clearly related to quality, while others are not; (3) the *directions* of relationships between policy variables and measures of quality; and (4) perhaps most important, the *consistency* of findings--the fact that the relatively independent sets of quality measures--caregiver behavior, child behavior and test scores--all point to the importance of the same policy variables--classroom composition (especially group size) and child related education/training.

References

Abt Associates, *National Day Day Care Study First Annual Report: Overview of the Study and Phase II Design*. Volumes I and II. Cambridge, Mass.: Abt Associates Inc., 1976.

Bache, W.L. "Comparing Alternative Measures of Classroom Composition." *National Day Care Study Measurements and Methods*. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980a.

Bache, W.L. "A Psychometric of the National Day Care Study Phase III Child Test Battery," *National Day Care Study Measurements and Methods*. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980b.

Cronbach, L.J., Gleser, G.C., Nanda, H., and Rejartnam, N., *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles."* New York: Wiley, 1972.

Goodrich, N.N. "The Effects of Day Care in Eight Atlanta Public Schools Day Care Centers." *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980b.

Goodrich, R.L. and Singer, J.D., "Cognitive Change in the NDCS", *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Goodson, B.D. "The Adult-Focus Observation Effects Analysis," *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Medley, D.M., and Mitzel, H.E., "Measuring Classroom Behavior by Systemztic Observation," in *Handbook of Research on Teaching*, N.L. Gage (Ed.), Chicago, Ill.: Rand McNally, 1963.

Mosteller, F., and Tukey, J.W. *Data Analysis and Regression*. Reading, Mass.: Addison-Wesley, 1977.

Prescott, E., Jones, E., Critchensky, S., Milich, C., and Hazelhoef, E., "Assessment of Child Rearing Environments: An Ecological Approach," Pasadena, Calif.: Pacific Oaks, 1975.

Ruopp, R., Travers, J., Glantz, F. and Coelen, C., Children at the Center. Final Report of the National Day Care Study, Volume I. Cambridge, Mass.: Abt Associates Inc., 1979.

Singer, J.D. "Classroom Process-Child Outcome Analyses. National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Singer, J.D., Affholter, D.P., Goodrich, R.L., "Variance Components and Dependability of Measures Used in the National Day Care Study," Unpublished, 1978.

Stallings, J., and Broussard, P. Final Report. National Day Care Cost-Effects Study: Spring 1977 (Phase III) Data Collection. Menlo Park, Calif.: Stanford Research Institute, 1977.

Stallings, J., Wilcox, M., and Travers, J., "Phase II Instruments for the National Day Care Csot Effects Study: Instruments Selection and Field Testing," Menlo Park, Calif.: Stanford Research Institute, 1976.

Travers, J., Coelen, C., and Ruopp, R., National Day Care Study Second Annual Report: Phase III Design. Cambridge, Mass.: Abt Associates Inc., 1977.

Analysis of Test Score Growth in the National Day Care Study

Robert L. Goodrich

Judith D. Singer

TABLE OF CONTENTS

Page

INTRODUCTION AND ABSTRACT 212

CHAPTER ONE: MEASUREMENT OF COGNITIVE CHANGE 215

- Why Must We Measure Cognitive Change? 215
- Description of Tests 216
- Growth System Theory 217
- Analysis of the Longitudinal Subset of Children 222
  - Model Identification 227
  - Robust Estimates of Parameters 228
- PSI Measures of Change 230
  - Longitudinally Based Measure of Change 230
- Summary 236

CHAPTER TWO: ANALYSIS OF CHILD-LEVEL BACKGROUND VARIABLES 237

CHAPTER THREE: AGGREGATION AND THE UNIT OF ANALYSIS 244

- Introduction 244
- Mathematical Formulation 247
  - Hypothesis Tests at Child and Weighted-Center Levels 252
- Summary 255

CHAPTER FOUR: CALCULATING CENTER AND CHILD COMPONENTS OF VARIANCE 257

- Fixed Effect R^2 260
- Summary 263

CHAPTER FIVE: FALLIBILITY OF DEPENDENT AND INDEPENDENT MEASURES 267

- Summary 272

CHAPTER SIX: CENTER-LEVEL COGNITIVE EFFECTS ANALYSIS AND VALIDATION OF FINDINGS 273

- Description of Variables 273
  - Dependent Variables 273
  - Classroom Structure Variables 274
  - Caregiver Qualifications Variables 275
  - Covariables 275

236

TABLE OF CONTENTS (continued)

Page

Analytic Issues 276

Unit of Analysis 276
Sample Definition 277
Approach to Regression Model Construction 278

Data Analysis 279

Preschool Inventory (PSI) 279
Peabody Picture Vocabulary Test (PPVT) 294

Summary 299

Notes 306

References 308

Appendices 313

INTRODUCTION AND ABSTRACT

Two of the most important dependent measures in the National Day Care Study (NDCS) are the Preschool Inventory (PSI) and the Peabody Picture Vocabulary (PPVT) tests, each of which was administered in October, 1976 and May, 1977. These tests are regarded as measures of certain areas of cognitive development that may be accelerated or decelerated by environmental input. Bache (1980) discussed these cognitive areas and other psychometric considerations in a separate report. This report concentrates not on the psychometric meaning of the measures but rather on the analysis of the influence of day care characteristics on cognitive development. Because the dependent measures assess a complex, continuing, developmental process that occurs for all children whether or not they are enrolled in day care programs, and because there are many potential influences on the process, the analysis is complex. This report presents a step-by-step description of the cognitive effects analyses, and their methodological foundation. The following paragraphs summarize the content of each chapter.

(1) Measurement of Cognitive Change

The cognitive effects analysis focuses on cognitive change rather than on absolute level. The development of valid measures of change for the PSI and PPVT began with an analysis of the distribution of longitudinal test score growth curves. This statistical analysis led to the construction of "Generalized Change Scores" (GCS) that most accurately reflect influences such as that of the day care environment. The technique is closely related to Bryk's (1977) "value-added" analysis.

(2) Analysis of Child Level Background Effects on PSI and PPVT Change

Generalized Change Scores were used as dependent measures in child-level regressions against child background variables. It was determined that so little variance is predicted by background variables such as socioeconomic status that they can be virtually disregarded in analyses. Thus, complicated "covariable adjustment" strategies are unnecessary. A small adjustment for race of child was retained for analyses of the PPVT.

(3) The Unit of Analysis

The classroom is not a practical choice as unit of analysis because, in most cases, their enrollments are continually changing. Thus, the UOA must be the child or the center. It was demonstrated that statistically and practically significant intracenter correlation effects exist and are large enough to invalidate hypothesis testing at the child level. Therefore, the center was selected as the UOA. Each center was weighted by the number of children it contributed to the sample in order to minimize variance of statistical estimates and to obtain unbiased hypothesis tests.

(4) Child and Center Components of Variance

The relative importance of child and center level effects was assessed by calculating the child and center level variance components. From these the generalizability* of center means was calculated. It was shown that the center level component of variance of cognitive change is statistically and practically significant, and that center level means are of sufficient generalizability to permit useful center-level analyses. This further supports the choice of the center as unit of analysis.

---

*An extension of the notion of reliability. See Cronbach et al. (1972).

(5) Fallibility of Dependent and Independent Measures

An analysis of the effect of measurement error on the center-level regression model is presented; correction formulas are given for the required multivariate disattenuation correction.

(6) Center Level Cognitive Effects Analysis and Validation of Findings

Classroom organization variables (group size, staff/child ratio, number of teachers and aides) and caregiver qualifications were shown to be significantly associated with cognitive gains. Several effects models were constructed. Effects were validated by (1) cross-validation across segments of the sample, emphasizing the correspondence of results from a sample of eight centers from the Atlanta Public School System and those from the remaining centers; (2) robust regression techniques to reduce the influence of outliers on overall conclusions; (3) sensitivity analysis of weighting by center; (4) analysis of sensitivity to outlier centers; (5) analysis of sensitivity to center level measures; (6) comparison of results from the PSI and the PPVT.

CHAPTER ONE: MEASUREMENT OF COGNITIVE CHANGE

Why Must We Measure Cognitive Change?

The purpose of the Phase III cognitive effects analysis was to determine influence of center level day care variables (especially those regulatable by public policy) on children's cognitive development, controlling for external influences unrelated to the day care experience. The available data include (1) pretest and posttest scores on two cognitively oriented tests, (2) a number of center descriptive characteristics, and (3) observations of staff/child ratio, group size, numbers of caregivers, and measures of education and experience of observed caregivers. The observationally based data were collected during morning hours five times, at approximately six month intervals, between pretest (T3: October, 1976) to posttest (T4: May, 1977).* The tests measure two dimensions of the complex process of cognitive development. In order to separate day care effects from maturation and status (background) effects, the study must focus on the process of developmental change, an exceedingly complex and controversial methodological area (Cronbach and Furby, 1970).

Measures of changes were constructed by statistical analysis of entire growth trajectories and their longitudinal characteristics, in a formulation related to that of Bryk and Weisberg (1977), and Bryk (1977). These measures, termed generalized change scores (GCS), were used directly as dependent variables in effects analyses.

This choice of analytic method was not the only considered. In particular, we considered using the posttest score directly as the dependent measure and the pretest score as one of the independent variables. We rejected this approach for the following reasons.

---

*See Bache (1980a) for a description of these measures.

(1) The approach is not a dynamic one even though the process is. It applies traditional static analysis to a problem where change is the construct of direct psychological interest.

(2) By viewing the problem longitudinally it is possible to construct a measure of change that minimizes chances of bias due to improper correction for maturational differences across the study population. (See Bryk and Weisberg, 1977.)

(3) Test generalizabilities (Cronbach et al., 1972) can be accurately computed from longitudinal analysis of whole scores, without reliance on internal consistency criteria. Longitudinally derived measures of change require no additional correction for attentuation effects.

The following sections detail development of an appropriate measure of change.

Description of Tests

The two cognitive measures used in Phase III are the Preschool Inventory (PSI) test in the shorter 32 item version (Educational Testing Service, 1970), and the Peabody Picture Vocabulary Test (Dunn, 1959); Meissner, Shipan, and Gilbert, 1972). A few motor development items were also administered but did not yield a sufficiently reliable motor development scale. The PSI is a measure of certain school-relevant aspects of skill and general knowledge; it has been used in several previous studies, including Head Start Planned Variation (HSPV) and the Home Start evaluation (Love et al., 1976). The PPVT is a widely used measure of receptive language functioning. Although both the PSI and PPVT correlate highly with Stanford-Binet scores, they are not viewed as intelligence tests in the National Day Care Study. Both tests are useful measures of school readiness in regular classroom settings. A detailed psychometric analysis of the PPVT and the PSI in the NDCS and other studies is provided by Bache (1980b).

Growth System Theory

An important premise of the cognitive analysis is that the data can be thought of as longitudinal measures (with measurement error) on a set of growth trajectories that together comprise a "growth system." Each trajectory represents the growth of one particular individual from the population of interest. This premise was first stated clearly by Bryk and Weisberg (1977), and was further elucidated by Bryk (1977), Strenio (1977), and Goodrich (1978). It has led to the formulation of growth system theory, bringing together ideas in developmental psychology, ordinary statistics, and stochastic systems theory. It is the statistical distribution of whole trajectories, rather than single measurements, that is the object of study. To provide an orientation to this problem, we will first consider it under simplifying assumptions made originally by Bryk and Weisberg:

(1) Each individual growth curve (trajectory) is a smooth, deterministic function characteristic of the individual. Trajectories differ randomly from individual to individual. If the trajectory for an individual can be identified, the future growth of the individual can be predicted.

(2) Each curve is linear (perhaps in a transformed metric), so that it can be described by the level α and slope β at a reference age. Thus, if α and β are known, the entire curve--past, present, and future--is known.

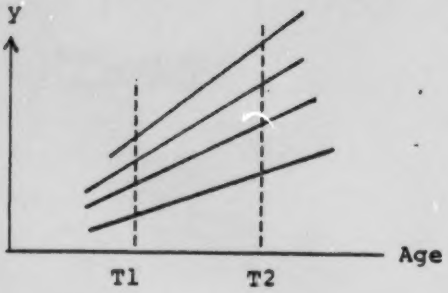
In this formulation, the growth characteristics of a population are described by the joint distribution of (α, β). Assuming normality, the distribution is determined by the mean $(\bar{\alpha}, \bar{\beta})$ and elements of the variance-covariance matrix Cov (α, β). It is common in evaluation theory to adjust for differences across populations by controlling for pretest scores. Bryk and Weisberg showed that, except under very special circumstances, this adjustment technique is biased.

Some of the distinct growth patterns that can occur are shown in Figure 1-4, adapted from Bryk and Weisberg. These growth systems can be used to demonstrate how adjustment bias arises.

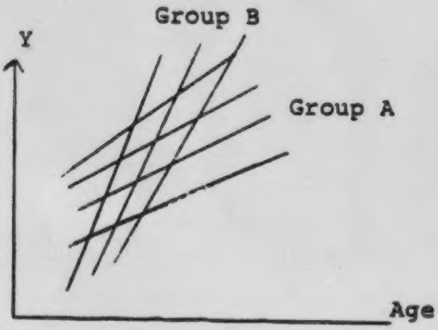
In Figure 1-1 (Fan Spread), the differences among individuals tend to increase with time, i.e., those who are further ahead initially tend to learn at faster rates. In Figure 1-2, (Fan Spread with Crossing), two groups are represented. Within group, differences tend to increase as in ordinary fan spread. However, the group that was initially behind overtakes and surpasses the initially more advanced group.

In Figure 1-3 (Fan Close), differences among individuals tend to decrease (and the curves would eventually cross). Fan close effects (but not crossing) can be caused by test ceiling effects. Figure 1-4 (Degenerate Fan Spread) is simply fan spread in which all individuals are initially equal.

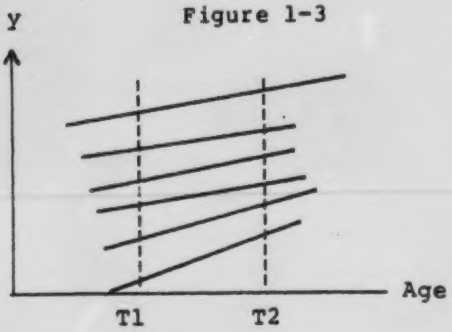
Fan Spread
Figure 1-1



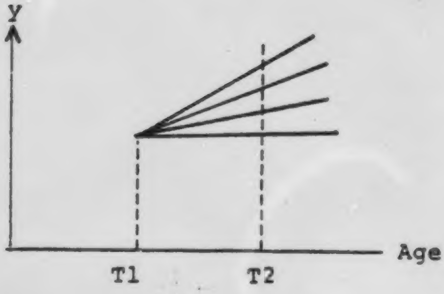
Fan Spread with Crossing
Figure 1-2


Group B
Y
Group A
Age

Fan Close
Figure 1-3



Degenerate Fan Spread
Figure 1-4



There are obviously many other patterns of growth that might occur--especially if trajectories are curvilinear. Nevertheless, these simple examples suffice to show that bias occurs in the most frequently applied adjustment methods. The magnitude of bias depends on the specific change measure used, and the nature of underlying differences in growth characteristics.

Kenny (1975) demonstrated that the results of an evaluation depend on the particular means used to construct the change measure. He considered four change measures (residual change, residual change with Lord-Porter correction, raw change, and standardized change) and showed that the null hypothesis has a different algebraic form in each case. Table 1.1 lists these null hypotheses for Kenny's four measures and for the general linear change score form. Each null hypothesis in Table 1.1 is stated in terms of the covariance of a function of population parameters and a treatment variable T, which may either be a dummy variable or an interval variable, such as group size or staff-child ratio. Assume that, in fact, there is no effect. Unless the populations are truly equivalent, at most one of the alternative null hypotheses can be true--since they all state different things. Unless the analyst has selected the "right" change measure, he or she will erroneously conclude (for large enough sample size) that an effect actually exists. The subject of this chapter is the identification of the "right" change measure through longitudinal analysis of the growth curves.

In practice, the situation is considerably more complex than that discussed so far--i.e., under the simplifying conditions cited earlier. For instance, the assumption of linearity of growth is usually not justified, at least for unscaled measures. The PSI test used in Phase II of the NDCS has only 32 items and begins to exhibit curvilinear

Table 1.1

NULL HYPOTHESES FOR ALTERNATE CHANGE MEASURES

Measure of Change		Null Hypothesis
Raw Change Score	$y_2 - y_1$	$\text{Cov}\left[(y_2 - y_1), T\right] = 0$
Standardized Change Score	$\frac{y_2}{\sigma_2} - \frac{y_1}{\sigma_1}$	$\text{Cov}\left[(y_2 - \frac{\sigma_2}{\sigma_2} y_1, T\right] = 0$
Residual Gain Score	$y_2 - \frac{r_{12\ 2}}{\sigma_1} y_1$	$\text{Cov}\left[(y_2 - \frac{\sigma_2 r_{12}}{\sigma_1} y_1), T\right] = 0$
Corrected Residual Gain Score	$y_2 - \frac{r_{12}\sigma_2 Y_1}{\sigma_1 \rho}$	$\text{Cov}\left[(y_2 - \frac{\sigma_2 r_{12}}{\rho \sigma_1} y_1), T\right] = 0$
General Form	$y_2 - By_1$	$\text{Cov}\left[(y_2 - By_1), T)\right] = 0$

y_1 = Pretest Score
σ_1 = Standard Dev. of y_1
y_2 = Posttest score
σ_2 = Standard Dev. of y_2
r_{12} = Correlation of y_1, y_2
ρ = Test Reliability
T = Treatment Variable

ceiling effects for children as young as 4 or 5. In addition, real growth trajectories include seemingly random increases or decreases associated with specific (but unpredictable) interaction events between child and learning environment, as discussed by McCall et al. (1973), and Goodrich (1978). The growth curves are, in other words, not usually smooth but "bumpy". The longitudinal analysis discussed in this report uses new statistical techniques (Goodrich, 1978) for the study of such complicated growth processes.

Summary. In even a very simple formulation of the growth curve model, the character of the model must be identified in order to select an unbiased measure of change. This fact has been discused by Bryk and Weisberg, Kenny, and Goodrich. Unfortunately, real human growth is even more complex than the linear, determinate models assumed in this section. The next section extends the models to provide a more realistic description of the growth process.

Analysis of the Longitudinal Subset of Children

Since the Phase III analysis involves only a pre- and a post-test, measures of change must be restricted to simple alternatives like straight change scores, residual change scores, standardized change scores, fraction of items learned etc. Potentially better measures of change can be constructed from longer sequences of longitudinal data. Because of this, we are interested in the subset of 110 children for whom four PSI tests are available (those children that participated both in Phase II and Phase III). For this longitudinal subject the entire 3-4 year age range is spanned. Unfortunately, the sample is too small and scattered across centers to serve as a basis for study of center level effects. However, it was valuable as a resource to study characteristics of the overall frowth process, and thus to construct a "generalized change score" (GCS) for use in analyses of center effects based on the entire sample.

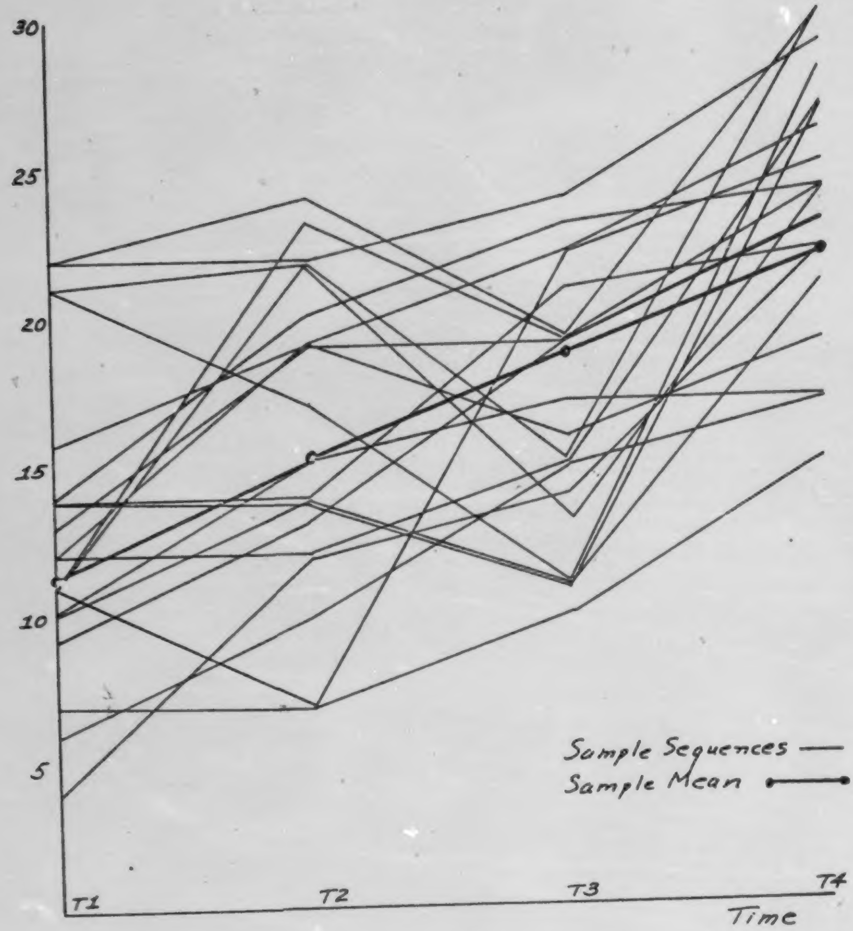
Figure 1-5 illustrates the stochastic nature of these data, which may be represented as a set of sequences $\{Z_{k,i};\ 1 \leq k \leq T,\ 1 \leq i \leq N\}$ where the cross-sectional sample size N is 110, and the longitudinal sample size is 4 (at times T1-T4). It is assumed that the i'th child's measurements $\{Z_{k,i};\ 1 \leq k \leq T\}$ are jointly described by the statistical distribution across the study population of parameters that describe each curve initially, a stochastic difference equation model of growth, and a measurement error model from classical test theory. The parameters of this model were identified by a maximum likelihood technique that employs the recursive Kalman filter equations (Kalman, 1960, Kalman and Bucy, 1961) to construct the likelihood function and its derivatives. These are used in conjunction with Newton-Raphson numerical maximization in a computer algorithm programmed in FORTRAN for the CDC 6400. A complete description of the model and the identification technique was provided by Goodrich (1978), and by Goodrich and Caines (1979a-c).

Three alternative stochastic models were considered as descriptors of the PSI trajectories. More complex models could not be considered because of the extreme shortness of the series available.

Model I is a first order autoregressive model with measurement error and deterministic constant inputs. The "true" PSI score $X_{k,i}$ is assumed to evolve according to the stochastic difference equation

$$X_{k+1,i} = A\, X_{k,i} + W_{k,i} + U_k \qquad \begin{matrix} 1 \leq k \leq 3 \\ 1 \leq i \leq 110 \end{matrix} \qquad (1\text{-}1)$$

Figure 1-5. **Sample Mean and Twenty Randomly Selected Data Sequences**



where $\{W_{k,i}\}$ are uncorrelated zero-mean random variables with variance σ_w^2 and $\{U_k\}$ is a sequence of unknown constants. The distribution of $X_{1,i}$ is assumed to be normal with mean $\bar{x}_{1/0}$ and variance $V_{1/0}$. Unknown parameters in (1-1) include the autoregressive constant A, the variance σ_w^2 of $W_{k,i}$ and the parameters $\{U_k\}$.

Each measurement $Z_{k,i}$ is related to the corresponding true score $X_{k,i}$ via the classical test theory measurement equation:

$$Z_{k,i} = X_{k,i} + V_{k,i} \tag{1-2}$$

In this equation, the measurement errors $V_{k,i}$ are assumed to be mutually uncorrelated and to have the common variance σ_v^2. The generalizability ρ, which will depend on the study population as well as the test, can be computed from σ_v^2 for any desired population, using the equation:

$$\rho = 1 - \frac{\sigma_v^2}{\mathrm{Var}(Z_{k,i})} \tag{1-3}$$

The parameters to be identified include:

$A,\ \sigma_w^2,\ \{U_k\},\ \bar{x}_{1/0},\ V_{1/0},\ \sigma_v^2.$

These make up a vector that will be denoted by θ. Model I appears deceptively simple because of its linear difference equation form. In fact, the equation describes a vast family complex of curvilinear trajectories with complicated longitudinal stochastic behavior.

Model II extends Model I to second order by allowing X to be a 2-dimensional vector--i.e., the model is of "second order". The equation corresponding to 1-1 is as follows:

$$X_{k+1,i} = \begin{bmatrix} A_1 & A_2 \\ 1 & 0 \end{bmatrix} X_{k,i} + \begin{bmatrix} G_2 \\ 1 \end{bmatrix} W_{k,1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} U_k \qquad (1-4)$$

$$Z_{k,i} = (0\ 1)\ X_{k,i} + V_{k,i}$$

Model II is a mixed, autoregressive moving average model, described by the parameters (A_1, A_2, G_2). The As are autoregressive parameters while G_2 is an "impulse response" parameter functionally related to the As and to a moving average parameter. Box and Jenkins label the model ARMA (1,2). The state X_k is a 2x1 vector whose second component represents the true PSI score; the initial state X_1 is assumed normal with vector mean $\bar{X}_{1/0}$ and covariance matrix

$$V_{1/0} = \begin{bmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{bmatrix}.$$ The processes $\{W_{k,i}\}$ and $\{V_{k,i}\}$

correspond to those described earlier. The parameter vector is now

$$\theta = (A_1, A_2, G_2, \sigma^2_w, \sigma^2_v, U_k, \bar{X}_{1/0}, V_{1/0}).$$

The increased number of elements of the parameter vector makes Model II harder to identify accurately than Model I.

Model III extends Model I in a different way. While Model I assumes that all subjects have equal mean growth rates, Model III includes subject-to-subject differences represented by the Term R in the following equations.

$$\begin{aligned} X_{k+1,i} &= A\,X_{k,i} + W_{k,i} + R_i + U_k \\ Z_{k,i} &= X_{k,i} + V_{k,i} \end{aligned} \tag{1-5}$$

The term R_i is randomly distributed across the population but is constant within case. We assume that $X_{1,i}$ and R_i have a joint normal distribution with mean $(\bar{X}_{1/0}, 0)$ and covariance matrix $\begin{bmatrix} \bar{V}_{1/0} & D_R \\ D_R' & \sigma_R^2 \end{bmatrix}$, so that the parameter set for the model is $\theta = (A, \sigma_w^2, U_k, \bar{X}_{1/0}, V_{1/0}, D_R, \sigma_w^2, \sigma_V^2)$. Model I can be derived from Model III by setting $D_R = \sigma_R^2 = 0$. Thus, Model I can be regarded as the mathematical expression of the null hypothesis $D_R = \sigma_R^2 = 0$.

The flexibility of the stochastic formulation can be verified by noting that it includes all deterministic linear models discussed by Bryk and Weisberg (see above). If one sets $A=1$, $\sigma_w^2 = 0$, $U_k = \text{const} = \bar{B}$, $\bar{X}_{1/0} = \alpha$, $V_{1/0} = \sigma_\alpha^2$, $\sigma_R^2 = \sigma_B^2$, $D_R = \sigma\ \alpha\ \sigma\ \beta\ r\ \alpha\ \beta$, to obtain the formulation described in Section 3.1 above. Model III extends the Bryk and Weisberg formulation to include both curvilinear mean growth (when $A \neq 1$) and stochastic (non-deterministic) fluctuations (when $\sigma_w^2 > 0$).

Model Identification

The longitudinal (Phase II-Phase III) data were analyzed by carrying out the following steps.

(1) All data were corrected to zero cross-sectional mean for each of the four time points. These reduced data were analyzed in a series of computer runs that tested the fit of each of the models I-III. On the basis of asymptotic likelihood ratio tests, the structure was determined to that of Model I. Model II was rejected at the level .10 and Model III

at the level .08. Parameter estimates and their standard errors were determined for the parameters A, σ_w^2, $V_{1/0}$, σ_v^2) that occur in Model I.

(2) The cross sectional means (subtracted out for Step 1)were analyzed to determine the constants (U_1, U_2, U_3, $X_{1/0}$) assuming that (A, σ_w^2, $V_{1/0}$, σ_v^2) had the values estimated in step (1), and using the technique described by Goodrich (1978). The null hypothesis $U_1 = U_2 = U_3$ was not rejected (chi-square test; $P = .13$). Therefore it was assumed that these constants were equal for purposes of parameter estimation. The resulting parameter estimates are given in Table 1-3.

(3) The model was checked and validated by verifying that residuals were statistically uncorrelated, that the predicted covariance matrix was not significantly different from the sample covariance matrix, and that the predicted mean curve achieved proper goodness-of-fit to the observed mean growth curve. Results of the latter test are exhibited in Table 1-4.

Robust Estimates of Parameters

Although ML estimates are asymptotically efficient for normally distributed data, they may not be robust with respect to severe departures from normality, especially when sample sizes are small. Recently much attention has been paid to so-called robust statistics, especially the jackknife method developed by Tukey (1978) and explained in detail by Mosteller and Tukey (1978). In this technique, the data are divided into several groups. Then separate parameter estimates are made for each reduced data set obtained by excluding one data group. Then, "pseudo-values" Y_j^* for parameter estimates are calculated by the equation:

$$Y_j^* = KY_o - (K-1)Y_j, \quad j = 1,2, \ldots,K \qquad (1\text{-}6)$$

where K is the number of groups, Y_o the parameter estimate

Table 1.2

SUMMARY OF STATISTICAL DATA

Model	Conditioning Of Hessian	LLF	Number of Parameters	Level of Significance
Theoretical Lower Bound	-	794.32	10	.14
Model I	Good	799.20	4	null hypothesis
Model II	Poor	795.42[a]	8	.10
Model III	Poor	796.7	6	.08

[a] Upper bound. Full minimum not achieved.

Table 1.3

PARAMETER ESTIMATES AND THEIR STANDARD ERRORS

Parameter	ML Estimate	Estimated Standard Error
A	.909	.05
$V_{1/10}$	17.38	3.07
σ_w^2	3.95	1.52
σ_v^2	5.18	1.10
$X_{1/0}$	11.30	-
$U_1 = U_2 = U_3$	4.89	-

Table 1.4

PREDICTED MEAN AND SAMPLE MEAN

	T1	T2	T3	T4
Sample Mean	11.37	14.89	18.16	21.96
Predicted Mean	11.30	15.16	18.67	21.87

using all available data, and Y_j that obtained for the estimate when the j'th group was omitted. The jackknife estimate and its variance are given by:

$$\bar{Y}^{*} = \frac{1}{K} \sum_{j} Y^{*}_{j} \qquad S^{*2} = \frac{1}{K} \; S^{2}, \qquad (1\text{-}7)$$

where S^2 is the sample variance of the pseudo-values.

In the present application, eleven groups of 10 subjects each were formed randomly. Separate maximum likelihood (ML) parameter estimates were made for each of the 11 resulting reduced data sets (see Table 1-5). Overall ML estimates and conventional parameter error estimates turned out to be close to jackknifed values. Thus, simple ML estimates appear to be reassuringly robust for these data.

Summary. A subset of the subjects were tested four times in the course of Phase II and Phase III. This subset was used to determine an underlying growth model that adequately describes both mean growth and subject-to-subject differences. The simplest model, denoted ARMA(1,0) with measurement error by Box and Jenkins, was selected by statistical hypothesis testing. Model parameters were determined and shown to be robust. The next section shows how a measure of change was derived from the identification of the underlying growth model.

PSI Measures of Change

Longitudinally Based Measure of Change

It was determined that the distribution of PSI developmental trajectory data is described by the following stochastic difference equation model:

Table 1-5

ESTIMATES OBTAINED BY JACKKNIFE METHOD

ML Estimates	A	σ_w^2	σ_v^2	$V_{1/0}$
Entire Sample	.90949	3.94711	5.17759	17.37589
1	.90686	3.57428	5.14253	18.16757
2	.90547	4.56481	4.77656	19.00743
3	.90223	3.78566	5.31813	17.90313
4	.90802	4.22483	5.05623	16.90015
5	.91361	4.11079	5.07464	17.33767
6	.90841	4.07115	5.15885	17.03517
7	.92226	3.00405	5.74061	17.40330
8	.89182	5.28138	4.53196	16.43901
9	.90115	4.54255	5.05568	16.89951
10	.90642	3.73532	5.33622	17.14301
11	.93534	2.69613	5.59384	16.40833
Jackknife Estimate	.91249	3.79007	5.28226	17.8218
SD of Jackknife Estimate	.0348	2.414	1.035	2.332
Conventional Estimate of SD	.0519	1.522	1.099	3.071

$$\begin{aligned} X_{k+1,i} &= A\, X_{k,i} + W_{k,i} + U \\ Z_{k,i} &= X_{k,i} + V_k \end{aligned} \qquad (1\text{-}8)$$

where

$$\hat{A} = .91 \quad \hat{\sigma}^2_w = 3.95 \quad \hat{\sigma}^2_v = 5.18$$

$$\hat{U} = 4.89 \quad \bar{X}_{1,i} = 11.3$$

$$V_{1/0} = \text{Var } X_{1,i} = 17.4$$

This model is the best fit to the 110 longitudinal records that does not involve any center level explanatory variables (e.g., staff-child ratio). In this section we develop a measure of change based on Model I. This measure of change will then be used to evaluate the effects of center level variables.

The principle to be used is a modification of the "value-added" principle originated by Bryk and Weisberg (see Bryk, 1977). First, a forecast of the posttest score based on all available prior data and known trajectory statistics is computed. Then the "value added" is computed as the difference of the actual and forecasted posttest score. This difference between optimal forecast and actual measure can be attributed, in part, to events that occur between pre- and posttests--i.e., to the day care environment. The method presented by Bryk and Weisberg and the method presented here differ principally in the technique that is used to prepare the forecast posttest score. While Bryk and Weisberg use cross-sectional data to calculate a mean growth curve (and hence estimated growth increments at any age), we use principles borrowed from time series theory to prepare an optimal forecast based on the underlying growth model.

It is well known (see, e.g., Astrom (1970)) that the optimal (i.e., minimum variance) forecast of $Z_{2,i}$ based on the previous measure $Z_{1,i}$ is prepared as follows. First, a minimum variance estimate of the true score $X_{1,i}$ is calculated by the formula

$$\hat{X}_{1,i} = \left(\frac{\bar{X}_{1/0}}{V_{1/0}} + \frac{Z_{1,i}}{\sigma_v^2}\right) / \left(\frac{1}{V_{1/0}} + \frac{1}{\sigma_v^2}\right) \qquad (1\text{-}10)$$

The right hand side of this equation is a weighted average of the measured score and the population mean score. Second, the forecast value of $Z_{2,i}$ is calculated by the equation

$$\hat{Z}_{2/1,i} = A\ Z_{1,i} + U, \qquad (1\text{-}11)$$

so that the value added is given by the formula:

$$\triangle' = Z_{2,i} - A\ \frac{(\sigma_v^2\ \bar{X}_{1/0} + V_{1/0} Z_{1,i})}{(\sigma_v^2 + V_{1/0})} \qquad (1\text{-}12)$$

The relative weights from equation 1-12 were determined to be given by:

$$\frac{\sigma_v^2}{\sigma_v^2 + V_{1/0}} = .23 \qquad \frac{V_{1/0}}{\sigma_v^2 + V_{1/0}} = .77 \qquad (1\text{-}13)$$

In the present applications, formula (1-12) was first replaced by the simpler form

$$\triangle_i = Z_{2,i} - A\ Z_{1,i} + \text{constant} \qquad (1\text{-}14)$$

This still produces an unbiased estimate of the effect $W_{k,i}$ given the pretest score. The result is an unbiased, near optimal calculation for value-added that is termed the "generalized change score". The constant is irrelevant to effects calculations, and can be omitted without loss, to obtain the form

$$\Delta_i = Z_{z,i} - .91\, Z_{1,i}.$$

Since the data to be analyzed are from Phase II, $Z_{z,i}$ is the test score at T4 and $Z_{1,i}$ is that from T3.

It thus turns out that Model I has led to the identification of the residual change score, disattenuated for reliability effects, as the appropriate change score to use for the PSI data. Normally one obtains the disattenuated residual change score by first regressing posttest scores against pretest, and then dividing the regression coefficient by the test reliability. Suppose that the regression equation so obtained is

$$Z_{2,i} = B\, Z_{1,i} + C + e_i,$$

and that the test reliability is ρ. Then the disattenuated residual change score is

$$Z_{2,i} - \frac{B}{\rho} Z_{1,i},$$

the same form determined through our longitudinal analysis. In fact, using $\rho = .85$ (Cronbach's alpha for the PSI) it was determined that

$$\frac{B}{\rho} = .86.$$

This is a remarkable convergence with the value .91 obtained from the longitudinal analysis. Note that if the best fitting model had been more complex than Model I, the generalized change score would have contained more terms and the similarity to the disattenuated residual change score would be less close.

The coefficient A for the model of the form given by (1-8) can be estimated by yet a third method if age-test covariances are known. Since $W_{k,i}$ is assumed to be uncorrelated with age, we find that

$$\mathrm{Cov}\ (Y_{k+1,i},\ a_i) = A\ \mathrm{Cov}\ (Y_{k,i},\ a_i)$$

where a_i is the age of subject i. Therefore, we can estimate A via

$$\hat{A} = \frac{\mathrm{Cov}\ (T4,a)}{\mathrm{Cov}\ (T3,a)} \qquad (1\text{-}15)$$

from T3, T4 covariances with age. Use of the formula yields A = .88 based on the Phase III sample of approximately 500 children.

If change-measures are not statistically independent of age then aggregation to center level would result in measures whose expected value depended on the age distribution in the center. It would be necessary to use age as an independent variable in analyses. However, for any of the suggested estimation techniques, relatively age-insensitive measures result. This is another verification that the model fits the growth curves very well over their entire range. Age, in this formulation, is merely an indexing variable with no causal role in the analysis.

Table 1-6 summarizes results obtained by the three different methods. The convergence exhibited in the table builds confidence that the PSI growth curve system has been properly identified.

Table 1-6

SUMMARY OF RESULTS FROM DIFFERENT ESTIMATION METHODS

	N	Estimate of B		Estimate of ρ
Longitudinal	110	91±.05	Longitudinal	.87
Regression with L-P Correction	800-900	.86	Cronbach's α	.85
Age-Covariance Method	800-900	.88		

Summary

The generalized change score (GCS) is the difference between the observed posttest score and the optimal predicted posttest score based on early information. A simplified form of the generalized change score was derived for use in analyses of center effects. Because of the simplicity of the underlying growth dynamics the GCS reduces to the disattenuated residual change score. Three different methods of calculation of the GCS lead to approximately the same result:

$$\triangle = z_{z,i} - .86\ z_{1,i}$$

This form was used for analyses of center effects, described in the next section.

CHAPTER TWO: ANALYSIS OF CHILD LEVEL BACKGROUND VARIABLES

In many cases the effects determined in quasi-experimental studies depend on the particular techniques selected to adjust for nonequivalence of treatment and comparison groups. For instance, the Head Start evaluation shows negative Head Start effects using ANCOVA and positive effects using Magidson's (1977) "true score" model. A vast literature on the problem exists, but, unfortunately, there is no consensus on the best methods to employ. In the current study, reliance on ANCOVA or other statistical adjustments was minimized by (1) attention to the problem in the design phase, i.e., by balancing race, social class, and other status measures whenever possible and (2) choice of the GCS method for evaluation of cognitive growth.

In this section we will investigate effects of background variables on generalized cognitive change scores at the child level. Table 2-1 lists the background variables considered in the analysis, including child-, neighborhood- and center-level variables. Center-level variables were obtained by aggregating over child-level variables, while neighborhood variables are from 1970 census tract data. Only child-level variables are considered in this section, since interest centered on the effect of background variables at the child level.

Missing data were a problem in the analysis. Only 489 cases included required test data and the full list of child-level covariables. These 489 cases were analyzed in stepwise regression runs (change measure against covariables); results are displayed in Tables 2-2 to 2-3. The variables most affected by missing data are the previous child care variables CNTRCARE and NOTCCARE. These account for for almost no variance in either GPSI or GPPVT. When CNTRCARE and NOTCCARE were excluded, sample size rose to N = 687. Results are shown in Tables 2-2B and 2-3B.

Table 2-1

BACKGROUND VARIABLES ON DATA FILES

Child Level

Computer Designation	Definition
RACE*	Race 1=B 2=W
SEX*	Sex 1=M 2=F
AGE*	Age as of January 1, 1977
TINCENT	Time in center as of January 1, 1977
NFOLKS*	Number of people at home
NADULT*	Number adults at home
NLT12*	Number of children less than 12 at home
NSIBLING*	Number of siblings
INCOMEC*	Income scale
MOTHRED*	Mother's education (years of education)
CNTRCARE*	Months of previous center care
NOTCCARE*	Months of preschool non-center child care

Neighborhood Level

CENSUS33	Percent of Families Feb., 1969 income was less than the poverty level
CENSUS03	Persons per household
CENSUS21	Females in labor force
TRACT9	Fraction pop. in same house as 1965

Center Level

FRWHITE	Fraction white (from RACE)
FRMALE	Fraction male (from SEX)
AGEMEAN	Mean age (from AGE)
AGEMAX	Maximum age in center
AGEMIN	Minimum age in center
NPEOPLE	Mean persons at home (from NFOLKS)
NADULTS	Mean adults at home (from NADULT)
NLT12H	Mean children less than 12 at home (from NLT12)
NSIBS	Mean siblings (from NSIBLING)
INCOME	Mean family income (scaled) (from INCOMEC)
CENTER	Mean months previous center care (from CNTRCARE)
NOTCNTR	Mean months previous non-center care (from NOTCCARE)

*Also present on Center Aggregated File

Table 2-2A

RESULTS OF STEPWISE REGRESSIONS, CHILD LEVEL, GPSI ON CHILD BACKGROUND VARIABLES

n=489

	Variable				R Square		
Step	Entered	Removed	F to Enter or Remove	Significance	R Square	Change	Simple R
1	NFOLKS		5.35312	.021	.01087	.01087	.10427
2	NADULT		1.63924	.201	.01420	.00333	.09760
3	INCOMEC		1.87832	.171	.01800	.00380	-.02121
4	AGE		1.33499	.248	.02070	.00270	.05228
5	CNTRCARE		1.17594	.279	.02308	.00238	-.04881
6	TINCENT		.75486	.386	.02461	.00153	-.03342
7	SEX		.17244	.678	.02496	.00035	.01983
8	MOTHRED		.16131	.688	.02528	.00033	-.01208
9	NSIBLING		.12314	.726	.02553	.00025	.05170
10	RACE		.09678	.756	.02573	.00020	-.00201
11	NOTCCARE		.07763	.781	.02589	.00016	.00206
12	NLT12		.01449	.904	.02692	.00003	.05773

Table 2-2B

GPSI AGAINST 10 REGRESSORS
(CNTRCARE, NOTCCARE EXCLUDED)

n=687

	Variable				R Square		
Step	Entered	Removed	F to Enter or Remove	Significance	R Square	Change	Simple R
1	NFOLKS		2.01111	.257	.00293	.00293	.04510
2	INCOMEC		.78145	.477	.00407	.00114	-.03279
3	NADULT		.60564	.437	.00495	.00088	.03521
4	RACE		.26220	.609	.00566	.00023	-.03340
5	AGE		.22325	.637	.00566	.00003	-.01682
6	SEX		.13959	.709	.00586	.00020	-.01303
7	MOTHRED		.08839	.766	.00599	.00013	-.01462
8	NSIBLING		.04807	.827	.00606	.00007	-.03331
9	NLT12		.04695	.829	.00613	.00007	.03676
10	TINCENT		.04209	.838	.00619	.00006	-.00759

Table 2-3A

RESULTS OF STEPWISE REGRESSIONS, CHILD LEVEL, GPPVT ON CHILD BACKGROUND VARIABLES

n=489

(Two variables were not entered by SPSS.)

	Variable				R Square		
Step	Entered	Removed	F to Enter or Remove	Significance	R Square	Change	Simple R
1	NFOLKS		5.35312	.021	.01087	.01087	.10427
2	NADULT		1.63924	.201	.01420	.00333	.09760
3	INCOMEC		1.87832	.171	.01800	.00380	-.02121
4	AGE		1.33499	.248	.02070	.00270	.05228
5	CNTRCARE		1.17594	.279	.02308	.00238	-.04881
6	TINCENT		.75486	.386	.02461	.00153	-.03342
7	SEX		.17244	.678	.02496	.00035	.01983
8	MOTHRED		.16131	.688	.02528	.00033	-.01208
9	NSIBLING		.12314	.726	.02553	.00025	.05170
10	RACE		.09678	.756	.02573	.00020	-.00201
11	NOTCCARE		.07763	.781	.02589	.00016	.00206
12	NLT12		.01449	.904	.02692	.00003	.05773

Table 2-3B

GPPVT AGAINST 10 REGRESSORS

n=687

(One variable was not entered by SPSS.)

	Variable				R Square		
Step	Entered	Removed	F to Enter or Remove	Significance	R Square	Change	Simple R
1	RACE		10.23584	.001	.01472	.01472	.12134
2	NADULT		5.06676	.025	.02197	.00724	.09617
3	NFOLKS		1.12892	.288	.02358	.00161	-.00655
4	NSIBLING		1.39050	.239	.25570	.00199	-.00438
5	TINCENT		.63652	.361	.02676	.00120	.03190
6	INCOMEC		.45946	.498	.02742	.00066	.06136
7	MOTHRED		.11479	.735	.02759	.00066	.06136
8	NLT12		.04062	.840	.02764	.00006	-.05374
9	AGE		.01029	.919	.02766	.00001	-.00082

Results of the multiple regression showed that for the PSI only .006 of the variance is accounted for by 10 covariables, and none of them are significant. For the PPVT, 10 variables account for .028 of the variance. There are two significant regressors: RACE and NADULT, accounting together for .02 of the variance. It was decided to retain RACE as a covariable but not NADULT. The latter accounted for only .007 of the variance which is not of practical significance. It was determined (see below) that the center-level effect is much larger than can be accounted for by any child-level covariables. Thus, there is little sensitivity of any of the conclusions to the details of child level background variable adjustments.

CHAPTER THREE: AGGREGATION AND THE UNIT OF ANALYSIS*

Introduction

Choice of the appropriate unit of analysis is a key decision in any investigation involving hierarchical levels of data. In educational evaluation, data are usually collected in an explicit hierarchy; children are nested within classes, classes within schools, schools within school districts. While evaluators want to focus on the response of individual children to educational treatments, the treatments are usually applied at class, school, or school district level. The question of individual response to treatment is often reformulated in the question of aggregate (class, school, or district) response. While the relation between individual and aggregate models at first appears to be straightforward, it is, in fact, extremely complicated. Decisions about the unit of analysis (UOA) are difficult and often affect the conclusions. Answers to specific questions are quite likely to be different at different levels of analysis; conduct of multiple level analyses may, therefore, merely postpone the decision until the results have been seen. That approach should be rejected since too much subjective bias is likely to enter in such *a posteriori* decision making. This section presents some of the technical and practical considerations that entered the decision-making process in the National Day Care Study.

The problems of aggregation and disaggregation received only passing attention (Gehlke and Biehel, 1934; Yule and Kendall, 1950) until Robinson's (1950) classic paper on the "ecological fallacy". Robinson showed that not only the strength but the direction of relationships may differ at the individual and aggregate levels. Sociologists were cautioned against interpreting "ecological correlations", i.e., those at aggregate level, as measures of individual-level correlations. Subsequent debate, predominantly in economics, sociology, and political science, focused upon disaggregation, i.e., the use of aggegate data to draw inferences about individuals (see Goodman, 1953; 1959; and Duncan and Davis, 1953). Then in the early 1970s, the topic gained popularity in educational research with works by Hannan (1971), Hannan and Burstein (1974), and Cronbach (1976). However, as Cronbach (1976) and Knapp (1977) point out, no adequate solution has been reached for determining at what level (or levels) analysis should be carried out.

---

*This section is an adaptation of Singer and Goodrich (1979).

Since the literature offered no definitive approach, the research staff of the National Day Care Study (NDCS) considered the problem nearly from scratch. This paper outlines how the UOA was selected through the considerations of a combination of practical, substantive, and methodological issues. We will discuss a mathematical formulation of the problem where the effects of aggregation on regression coefficients and hypothesis testing can be presented, and conclude with a presentation of the decisions actually made and the justifications for them.

In the NDCS, data were collected at three levels: child, class/lead teacher and center.* Observations of child behavior, scores on two standardized tests (the Caldwell Pre-School Inventory and Peabody Picture Vocabulary Test) and information on family background characteristics were gathered at the child level. At the caregiver level, background information was collected and observations of behavior were made. Measures of classroom composition (group size, number of caregivers, and the resulting staff/

---

*Lead teacher and class level are virtually identical; with infrequent exceptions, only one lead teacher was asigned to a specific class.

child ratio) were calculated as averages of repeated observations in classrooms. Finally, data on auspices, program philosophy and orientation, and director qualifications were collected at the center level. A wide variety of analyses were to be conducted where, in each instance, the data had to be aggregated or disaggregated to the selected UOA. Thus, a class-level analysis required both disaggregation of center-level variables and aggregation of child level variables that entered the analysis. In this paper we discuss the choices made for three types of studies involving (1) Child Pre-School Inventory (PSI) and Peabody Picture Vocabulary Test (PPVT) growth, (2) caregiver behavior, and (3) child behavior as dependent variables. All three studies used regression (with covariables) as the principle analytic tool.

We did not consider aggregating PSI and PPVT data to the class level because the composition of specific classes within a given center often changed dramatically over the course of the day care year.* Limiting the analytic sample to children who remained in the same class from fall to spring would severely reduce the sample--by 25 percent. Moreover, the stability of the classroom unit varied significantly from site to site and center to center; as a result, the sample would be concentrated in more stable sites and centers. Hence, only two alternatives were considered for child test data--child and center levels. Caregiver behavior variables could be analyzed at person/class or center level. Child behavior could be studied at any of the three levels. The final decisions were to analyze caregiver behavior at the person/class level, child behavior at the class level and test scores at the center level. In the next section, the problem is set up mathematically and the reasoning behind the PSI/PPVT test score decision is presented.

---

*Reorganization of classes as a function of shifting enrollment is very common in day care, especially in for-profit centers which expand and contract with demand.

Mathematical Formulation

To simplify the mathematical development, we will consider only two levels which we will call the child (or individual) and the class. The results can be generalized to multiple-level data. First we state some easily demonstrated statistical facts about two-level (child and class) hierarchical regression models in order to demystify some arguments often made. These facts are:

- Class-level and child-level estimates of the magnitude of global class effects are identical provided the models are homologous (in a sense to be defined), and that class-level effects estimates are child-weighted. In order to be homologous, the child-level regression model must include aggregate "ecological" effects. All other variables must also occur in both models.

- Even if class-level models are not child-weighted, child- and class-level estimates of global class effects converge to the same (true) values as the number of classes tends to infinity, i.e., both forms of estimation are consistent. For reasonably large samples of classes, child and class models yield similar but not identical results. Serious discrepancies indicate that the models are not homologous, i.e., they are specified differently, often in some subtle way.

- Even when effects estimates are identical at the two levels, hypothesis tests yield different results when intraclass correlation effects exist, in which case, the individual-level test is biased, and class-level hypothesis tests should be conducted.

- There is no serious diminution in statistical power incurred by aggregation to class level despite loss of degrees of freedom. (The decresae in degrees of freedom in the denominator under aggregation is approximately compensated for by the decrease of error variance obtained by taking means.)

In multiple hierarchical models, these facts argue for aggregation to the highest level for which significant intra-level correlation effects exist. Effects sizes can be calculated at lower levels if hypothesis tests are not believed and models are homologous. We will prove the first three assertions in this paper.

Suppose that a study has been conducted to analyze the effect of a classroom variable on children's school performance. Then three types of variables exist. Child-level variables, (e.g., race or test score) are attached to specific children. Aggregate variables are classroom-level averages of child-level measures (e.g., fraction Black or average test score). Global (class-level) variables describe class level characteristics (e.g., educational treatment, class size, and staff/child ratio). Class-level analyses involve global variables and aggregate variables only. Child-level analyses can involve all three types of variables.

Let us adopt the following notation:

Y_{ij} = dependent child-level variables for child i within class j ($i = 1, 2, \ldots, n_j$; $j = 1, 2, \ldots, m$; $\sum_j n_j = N$),

$Y_{.j} = \frac{1}{n_j} \sum_i Y_{i,j}$ = associated aggregate dependent class-level variable,

Z_{ij} = matrix of independent child-level variables for child i within class j, including interactions as distinct components,

$Z_{.j} = \frac{1}{n_j} \sum_i Z_{ij}$ = matrix of aggregate variables, including child-level interaction components,

X_j = the (9x1) vector of global (class-level) variables for class j,

e_{ij} = random error associated with child i within class j, and

f_j = random error associated with class j.

With this notation, we assume the following "true" child-level linear model:

$$Y_{ij} = AZ_{ij} + BX_j + CZ_{.j} + e_{ij} + f_j \ , \qquad (3\text{-}1)$$

where sample means of all variables have been corrected to zero. Both e_{ij} and f_j are assumed to be mutually uncorrelated zero-mean random variables with variances σ^2 and ν^2 respectively. Model (1) is the most general linear model that can be based on independent variables of the types discussed above. Notice that both Z_{ij} and its aggregate Z_j are entered.

Two types of effects are included in the above model: (a) fixed effects (involving A, B and C) attributable to specific child- and class-level variables; and (b) random effects (e_{ij} and f_j), representing unpredictable effects at both levels. This model combines notions of ordinary regression (fixed effects) and variance components analysis (random effects), and is a version of the "mixed model" (see Graybill, 1961). Econometricians term such models "error components regressions," (see e.g., Arora, 1973). When $\nu^2 = 0$, the model reduces to an ordinary regression. The disaggregated aggregate variable Z_j is essential in the model; it represents class ecological effects on the individual child.

Using ordinary least squares to estimate the coefficients A, B and C in (3-1) is equivalent to minimizing the residual sum of squares:

$$RSS \overset{\Delta}{=} \sum_{i,j} (Y_{ij} - AZ_{ij} - BX_j - CZ_{.j})^2, \qquad (3\text{-}2)$$

which can be decomposed into child and class components:

$$RSS_{child} \overset{\Delta}{=} \sum_{i,j} [(Y_{ij} - Y_{.j}) - A(Z_{ij} - Z_{.j})]^2 \qquad (3\text{-}3)$$

$$\text{and } RSS_{class} \overset{\Delta}{=} \sum_{j} n_j[Y_{.j} - (A + C)Z_{.j} - BX_j]^2, \qquad (3\text{-}4)$$

where $RSS = RSS_{child} + RSS_{class}$.

This decomposition can be reparameterized by recognizing that the sum (A + C) in (3-4) is a free parameter, and hence may be set equal to a new independent free parameter D. Then (3-3) and (3-4) contain no common parameters.

Equation (3-1) can be directly aggregated to obtain the homologous (Hannan, 1971) class level model,

$$\text{where:} \qquad Y_{.j} = DZ_{.j} + BX_j + (e_{.j} + f_j) \qquad (3\text{-}5)$$

Intraclass correlations of residuals $(e_{ij} + f_j)$ present in (3-1) are removed by aggregation, i.e., random child effects and class effects become confounded. The child-weighted residual sum of squares for (3-5) is

$$\sum_j n_j[Y_{.j} - DZ_{.j} - BX_j]^2,$$

identical to RSS_{class} from the child model. Since RSS_{child} and RSS_{class} contain no common parameters, RSS is minimized by independently minimizing RSS_{child} and RSS_{class}. The resulting estimates $(\hat{A}, \hat{B}, \hat{D})$, can be used to calculate $\hat{C} = \hat{D} - \hat{A}$. The estimate B obtained from (3-2) is identical to that obtained by minimizing the child-weighted class-level residual sum-of-squares (3-4), proving the first assertion. If the class-level analysis is not child-weighted, child- and class-level estimates of B differ but, as the number of classes becomes larger, they converge to the same values.

If the "ecological term" $CZ_{.j}$ had been omitted from (3-2), then RSS_{child} and RSS_{class} would no longer be independent. Class and child level estimates of B would differ because of this difference in specification--ecological effects are present in the class-level model but not in the child-level model. This is because the effects of Z_{ij} and Z_j (child-level covariable and the associated ecological measure) are confounded at the class level. At the higher level, for example, it is impossible to separate the possibly differential effects attributable to the race of an individual child and the racial composition of the class. Although estimates for global class effects do not differ in child and weighted class-level analyses, child-level analyses are more informative about the process as a whole. However, coefficients for global variables, the primary focus of NDCS investigations, are estimated most economically, and with no difference in outcome, in weighted class-level analyses.

In contrast, hypothesis tests at child and class levels obtained from regression packages such as SPSS, BMD and SAS are quite different, even if class-level regressions are appropriately weighted. In fact, we will show that (a) child-level hypothesis tests for the significance of global variables are usually not appropriate, and (b) aggregate-level weights that permit valid hypothesis testing can be constructed.

Hypothesis Tests at Child and Weighted-Center Levels

Hypothesis tests for regression coefficients involve F-statistics constructed as ratios of scaled chi-square statistics. Implicit in these tests are possibly mistaken distributional assumptions. As a result, routinely constructed tests of the significance of coefficients often yield incorrect rejection regions because the sums-of-squares used are not distributed as chi-square, even when residuals have normal distributions.

It is easily shown that, subject to normality conditions, RSS_{child} is distributed as $\sigma^2 \chi^2_{n-m-p}$ where σ^2 is the child component of residual variance and p is the number of parameters in A, i.e. the number of distinct child-level covariables entered into the model. Let q be the number of components in (B,D). Then it can be shown (See appendix) that RSS_{class} is distributed as the sum of squares

$$\sum_{k=1}^{m-q} \lambda_k W_k^2, \qquad (3\text{-}6)$$

where the W_k, $1 \leq k \leq m-q$, are identically independently distributed (I.I.D.) normal random variables with zero mean and unit variance (N(0,1)), and λ_k, $1 \leq k \leq m-q$, are the non-zero eigenvalues of

$$\Sigma^{1/2} (I-H) \, \Sigma^{1/2} .$$

Σ is the diagonal matrix

$$\mathrm{diag}\,(\sigma^2 + n_1 \nu^2, \; \sigma^2 + n_2 \nu^2, \; \ldots, \; \sigma^2 + n_m \nu^2),$$

and H is the " Σ-weighted hat matrix"

$$\Sigma^{-1/2}\, X\, (X'\, \Sigma^{-1/2} X)^{-1}\, X'\, \Sigma^{-1/2}.$$

In this form X is the mxg matrix of global classroom variables assembled from $X_1, X_2, \ldots, X_m$. It can be seen that H has rank q and is idempotent. Therefore (1-H) is an idempotent mxm matrix of rank m-q .

In general, (3-6) does not have a chi-square distribution. Therefore, RSS_{class} and RSS are not in general distributed as chi-square. As a result, F-tests that depend upon chi-square distributional assumptions are invalid. To construct valid exact tests, it would be necessary to calculate the cumulative distribution of (3-6) numerically (Imhoff, 1961; Jensen and Solomon, 1972). However, σ^2 and ν^2 are not known and must be estimated to perform the required numerical integration.

The form (3-6) is chi-square if and only if all the $\lambda_k = 1$ or scaled chi-square if all the λ_k are constant. This occurs for either of the following special cases.

Condition A: $\nu^2 = 0 \implies RSS_{class} \sim \sigma^2 \chi^2_{m-q}$

Condition B: $N_1 = N_2 = \ldots = N_m = N/M \implies RSS_{class} \sim (\sigma^2 + \frac{N}{m} \nu^2) \chi^2_{m-q}$,

where m is the total number of classes. Thus, if there are no random class-level effects (Condition A), or if all classes are of equal size (Condition B), RSS_{class} is distributed as scaled chi-square.

In order to proceed most simply we will assume that condition B holds, i.e. that

$$RSS_{class} \sim (\sigma^2 + \frac{N}{m} \nu^2)\, X^2_{m-q},$$

Practically speaking this is a reasonable approximation whenever class sizes are not extremely uneven. However, as $\mathrm{RSS}_{\mathrm{class}}$ may be made chi-square even for unequal class sizes by appropriate weighting. This would ordinarily not be necessary. The more serious problem is the distribution of RSS. Even if $\mathrm{RSS}_{\mathrm{class}}$ and $\mathrm{RSS}_{\mathrm{child}}$ are separately chi-square, the mixture RSS = $\mathrm{RSS}_{\mathrm{class}} + \mathrm{RSS}_{\mathrm{child}}$ is not. In fact the distribution of

$$\mathrm{RSS} \quad \sigma^2 \chi^2_{n-m-p} + (\sigma^2 + \frac{N}{m}\nu^2)\ \chi^2_{m-q'} \quad , \qquad (3\text{-}7)$$

is not chi-square unless the class level random effects variance ν^2 is zero. We may not assume that

$$\mathrm{RS}\widetilde{\mathrm{S}} \quad \sigma^2 (\chi^2_{n-m-p} + \chi^2_{m-q}) = \sigma^2 \chi^2_{m-q}$$

unless the intraclass correlation   of class level residuals satisfies the domination relation

$$\frac{\rho}{1-\rho} << \frac{m}{N} ,$$

where N/m is the average number of children per class. Thus even small intraclass correlations make the chi-square assumption for RSS untenable when N/m is large.*

Child-level hypothesis tests are thus usually invalid. The conventional child-level F-test is biased (for $\nu^2 > 0$) in the sense that when the null hypothesis is true, it is nevertheless rejected with probability one as sample tends to infinity. These results are of practical consequence. In the NDCS, data were analyzed (experimentally) both at child and child-weighted center levels, using homologous models. Center-level effects estimates were identical up to round-off error, but many more effects were "significant" at child than at center level. The difference, mathematically and experimentally, results from too frequent rejection of the null hypothesis. It is not due to greater statistical power but to test bias.

---

*Even when $\nu^2 = 0$, problems may occur. For example, if there is specification error in the child-level model, Residual SS is not distributed as chi-square but rather as a mixture similar to that obtained when $\nu^2 \neq 0$. In this situation, hypothesis tests for B are also invalid.

The mathematical arguments cited above indicate that one should aggregate to the lowest level for which intraclass correlation effects cannot reasonably be assumed to vanish. In doing so one gains validity of hypothesis tests without losing substantial statistical power, but loses detail about effects that were aggregated over. Note, however, that the magnitude of child-level effects can be estimated even if hypothesis tests are invalid.

The next chapter describes the actual calculation of various components needed to apply the theory described above. It was determined in analyses of center-level (ν^2) and child-level (σ^2) components of variance of residuals that, in the notation given earlier, $\frac{N}{m}\nu^2$ and σ^2 were of the same order of magnitude, so that analyses at child level would yield biased hypothesis tests. Therefore, despite possible loss of child-level detail, the decision to conduct center level analysis was considered mandatory.

Summary

We have presented some of the considerations which entered into the process of selecting analytic units in the NDCS. Synthesis of this information led to the decision to analyze caregiver behavior at the person-level, child behavior at the class level and test scores at the center level.

The decision to analyze test scores at center level has already been supported. The lack of generalizability of staff-child ratio and group size at the child level provided further weight to the decision.

Child behavior could not be analyzed at the child level because results presented above showed that hypothesis tests were likely to be biased. Variance components analyses indicated that NDCS child behavior data were less well characterized as measures of child traits than as measures of overall classroom (or center) processes. The choice between class and center level was made on the basis of greater detail at class level, i.e. the ability to illuminate differences from class to class within center.

Caregiver behavior was investigated at the person level, which is essentially identical to class level since there was usually only one lead teacher per class. The greater generalizability of dependent and independent measures at the class/teacher level informed the decision to conduct teacher-level analyses.

These decisions were an important part of the analytic framework for NDCS analyses. Central to this approach, however, was the fact that the focus was on estimating the effects of staff-child ratio and group size on caregiver and child outcomes, not upon the relationships between person or aggregate-level variables and these dependent measures. If the emphasis was on estimating the relationship between a child's race, the ethnic composition of a class and child test scores, analyses might have been pursued at the child level, despite the technical difficulties cited.

CHAPTER FOUR: CALCULATING CENTER AND CHILD COMPONENTS OF VARIANCE

In the previous chapter, the problem of selecting a unit of analysis (level of aggregation) was related to determination of the relative magnitude of child- and center-level variance components. This chapter presents the analysis that was undertaken in order to determine these variance components. The primary analytical model is the random effects ANOVA (see Graybill's (1961) "Model V", p.351-354). The results were used to calculate the generalizability of center-level means, to assist in the interpretation of aggregated center-level models of day care effects, and to test for the appropriate unit of analysis, as discussed in Chapter Three.

The model presented in Chapter 3 (Equation (3-1)) is sometimes called an "error components regression" model because the conventional error term is split into two components representing child- and center-level errors. The analysis of the error components model is made clearer if we group certain terms from Equation 3-1. Recall that the grand mean has been subtracted out and that all quantities have zero means.

The terms with a single center index j are grouped together as

$$\alpha_j = BX_j + CZ_{.j} + f_j ,$$

representing the total center effect. The doubly indexed terms are grouped as

$$W_{ij} = AZ_{ij} + e_{ij} .$$

The resulting mathematical model is

$$Y_{ij} = \alpha_j + W_{ij} \qquad (4\text{-}1)$$

where

i = center index; $1 \leq j \leq m$
m = number of centers
i = child index, nested within center; $1 \leq i \leq n_j$
n_j = number of children in center j; $\Sigma n_j = N$
α_j = effect of center j
W_{ij} = effect of child i within center j, confounded with measurement error

In the analysis α_i and W_{ij} are independent random variables with zero mean and variances σ_α^2 and σ_ω^2 respectively. For purposes of hypothesis testing, all random variables are assumed to be normally distributed, but this assumption is not required for point estimations.

Table 4-1 gives the ANOVA table for this design. EMS estimators for σ_ω^2 and α^2 are obtained from the table by matching mean and expected mean squares. The explicit formulas are:

$$\sigma_\omega^2 = MS_k$$

$$(4\text{-}2)$$

$$\sigma_\alpha^2 = (MS_c - MS_k) \frac{n(m-1)}{n^2 - \Sigma n_i^2}$$

Table 4-1

ANOVA TABLE FOR CENTER/CHILD DESIGN

	DF	SS	MS	EMS
Total	N	$\Sigma_{ij} y_{ij}$		
Grand Mean	1	$\Sigma_{i,j} y^2_{..}$		
Centers	m-1	$SS_c = \Sigma_{i,j}(y_{.j} - y_{..})^2$	MS_c	$\sigma^2_\omega + \frac{n^2 - \Sigma n_j^2}{n(m-1)} \sigma^2_\alpha$
Children	n-m	$SS_k = \Sigma_{i,j}(y_{ij} - y_{.j})^2$	MS_k	σ^2_ω

This design allows simple calculation of the generalizability of center means. Since

$$E\,\Sigma_{i,j} (Y_{.j} - Y_{..})^2 = (m-1)\sigma^2_\omega + \frac{n^2 - \Sigma n_j^2}{n}\sigma^2_\alpha, \qquad (4\text{-}3)$$

the center level sum-of-squares includes variance attributable both to center effects (term involving σ^2_α) and to aggregated child-effects and error (term involving σ^2_ω). The fraction attributable to center effects is defined as the generalizability of the center mean and is given by

$$\rho = \frac{\frac{n^2 - \Sigma n_j^2}{n}\sigma^2_\alpha}{(m-1)\sigma^2_w + \frac{n^2 - \Sigma n_j^2}{n}\sigma^2_\alpha} \,. \qquad (4\text{-}4)$$

Thus an estimate of ρ is given by the simple formula

$$\hat{\rho} = \frac{MS_c - MS_k}{MS_c} \quad , \qquad (4\text{-}5)$$

where the terms on the right are defined in Table 4-1.

Fixed Effect R^2

Suppose that an ordinary (fixed effect) regression is conducted in which each center is represented by a separate (dummy) variable. The variance explained by such a regression is SS_c and the R^2 achieved is given by $SS_c/(SS_c + SS_k)$. The R^2 includes all variance explained by center sample means and thus includes capitalization on chance fluctuations of aggregated error and child effects. It is thus an over-estimate of the explanatory power of center-level causes. The analysis given so far, however, shows that the fraction

$$\rho \quad \frac{SS_c}{SS_c + SS_k} = (m-1) \, \frac{(MS_c - MS_k)}{SS_c + SS_k} \; , \qquad (4\text{-}6)$$

which we will call the "corrected R^2", is an estimate of the fraction of variance actually associated with center-level effects, while $(m - 1)\ (MS_c - MS_k)$ is an estimate of the total variance associated with center effects. The magnitude of the center effect in test score points is given by the square root of the center component of variance, i.e., by σ_α.

The calculations that have been discussed are made easily by performing an ordinary (fixed effect) one way ANOVA classification of the dependent variable by center. Results for several dependent measures are given in Tables 4-2 and 4-3.

Table 4-2

VARIANCE COMPONENTS OF GENERALIZED COGNITIVE CHANGE MEASURES

Measure	Component	DF	SS	MS	Estimated Root Variance Component
GPSI	Center	56	1886.5	33.69	$\sigma_\alpha = 1.14$
	Child	854	10771.9	12.61	$\sigma_\omega = 3.55$
GPPVT	Center	56	7857.9	140.32	$\sigma_\alpha = 2.30$
	Child	801	47592.1	59.41	$\sigma_\omega = 7.70$
RPPVT	Center	56	7028.6	125.51	$\sigma_\alpha = 2.18$
	Child	740	43246.0	58.44	$\sigma_\omega = 7.64$

Table 4-3

COMPUTED STATISTICS FROM VARIANCE COMPONENTS ANALYSIS

	Fixed Effect R^2	Corrected R^2	Significance	Generalizability of Center Mean
GPSI	.149	.093	.000	.63
GPPVT	.142	.082	.000	.58
RPPVT	.140	.075	.000	.53

The discussion in Chapter Three showed that the question of the appropriate UOA revolves about the variance components ν^2 and ν^2 of the error in the model (3-1). These components can be estimated by the approximate technique to be outlined here. Then, given values for $\hat{\sigma}^2$ and $\hat{\nu}^2$, the UOA can be selected.

The technique is as follows. It is known from regressions conducted at child and center levels (the latter are treated in Chapter Six) that to a reasonable approximation the coefficients A and C in Eq. (3-1) are zero, so that the model simplifies to

$$Y_{ij} = BX_j + e_{ij} + f_j$$

$$1 \leq i \leq n_j \ , \ 1 \leq j \leq m \ ,$$

where $E\ e_{ij}^2 = \sigma^2$ and $E\ f_j^2 = \nu^2$. The coefficient vector B in this model can be estimated consistently either at child or (weighted) center level. Then $\hat{\sigma}^2$ and $\hat{\nu}^2$ can be determined by variance components analysis of the residuals $Y_{ij} - \hat{B}X_j$. When the number of components in the vector X_j is small compared to the number of centers, there is no need to consider degrees of freedom used in estimating B, and the residuals $r_{ij} = Y_{ij} - BX_j$ can be analyzed directly, assuming that

$$r_{ij} = e_{ij} + f_j$$

The estimates for σ^2 and ν^2 include specification error variance at these two levels.

Estimates of σ^2 and ν^2 were obtained through one-way analyses of variance as described earlier in this chapter and as shown in Table 4-3A.

Table 4-3A

COMPONENTS OF RESIDUAL VARIANCE

Level	GPSI	RPPVT
Child (σ^2)	12.3	59.84
Center (ν^2)	.74	2.63

It was determined from these figures (see Appendix 1) that approximately optimal weights for center-level analyses are given by:

	Weight
GPSI	$(n_j^{-1} + .06)^{-1}$
RPPVT	$(n_j^{-1} + .04)^{-1}$

Since average n_j equals about 15, both terms of these formulas are important. Neither weighted nor unweighted center-level yield exactly appropriate hypothesis tests at center level, although neither would be in severe error. Child-level tests, however, are totally invalid since the residual sum-of-squares is not distributed as chi-square for any choice of weights. Departures from chi-square are severe for the values of σ^2 and ν^2 estimated.

Summary

The analysis shows that center effects must be considered large (1.14 points on the PSI, 2.18 on the PPVT), compared to "normal" growth. Mean growth (T4 - T3) was approximately 4.8 points for the PSI, 7.9 points for the PPVT, so that center effects can be reexpressed developmentally as 1.2 months (PSI) and 1.4 months (PPVT). In two other studies, mean development rate was slower than in the NDCS sample, viz.

(1) For the Home Start National Evaluation Sample, cross-sectional data suggest a growth rate of 2-2.5 points per 6 month period for the PSI 32 item test. This

figure does not include Home Start Treatment Effect since it depends entirely on pretest data. Source: High/Scope, 1973a (p. 123), 1973b (p. 148), 1974 (Table IV-47). See Table 4-4.

(ii) For the Head Start Planned Variation Study, similar estimates range from about 1 to 3 points per 6 month period. See Table 4-5.

Center effects as defined in this variance components study include all possible center level effects including aggregated child background effects. However, these were found to be negligible at the child level. They may still account for variance at the center level through so-called ecological effects. The total center-level variance is potentially only .63 (PSI) or .58 (PPVT) explainable through center-level regression model, through a computation of center-mean generalizability. Otherwise put, $(1.14)^2$ and $(2.18)^2$ are maximum center-level explainable variances. The fact that these figures are developmentally large shows that there are likely to be highly significant regression models at center level.

A variance components analysis of residuals obtained from center-level regression models (Chapter Six) shows that child level hypothesis tests are in serious error and weighted center-level tests are approximately correct.

Table 4-4

PSI32 MEANS AND STANDARD DEVIATIONS BY AGE GROUP: FALL 1972, SRPING 1973 AND FALL 1973 HOME START NATIONAL EVALUATION*

	Fall 1972			Spring 1973			Fall 1973		
Age (mos.)	N	Mean	SD	N	Mean	SD	N	Mean	SD
34–39	18	8.7	4.1	3	9.3	3.5	50	7.4	3.6
40–45	29	11.2	5.3	19	11.0	3.7	113	8.3	3.6
46–51	44	13.1	6.1	27	11.0	4.9	125	9.9	4.3
52–57	19	13.1	4.5	37	13.8	5.4	97	11.8	4.9
58–63	39	15.1	6.0	33	16.4	4.8	33	13.5	5.0
64–69	17	17.4	5.9	26	17.8	6.6	7	18.9	4.6
70–75				16	20.5	6.3			
				10	21.0	5.2			

*High/Scope, 1973a (p. 123), 1973b (p. 148), 1974 (Table IV-47)

Table 4-5

PSI32 MEANS AND STANDARD DEVIATIONS BY AGE GROUP AND PREVIOUS PRESCHOOL EXPERIENCE AND BY AGE GROUP AND RACE: FALL 1971 HEAD START PLANNED VARIATION STUDY

	Without Previous Preschool			With Previous Preschool		
Age (mos.)	N	Mean	SD	N	Mean	SD
36-38	4	7.8	4.82	--	--	--
39-41	4	6.8	1.48	--	--	--
42-44	16	7.6	3.79	2	12.0	3.00
45-47	63	10.2	4.65	6	13.0	3.27
48-50	207	10.6	4.51	35	12.1	5.86
51-53	374	11.2	4.93	63	13.0	5.31
54-56	397	12.5	5.02	57	13.2	5.60
57-59	368	13.4	5.11	81	15.0	5.59
60-62	257	15.9	5.56	121	17.7	5.61
63-65	162	17 0	5.40	99	17.2	5.81
66-68	165	17.4	5.59	96	19.7	5.09
69-71	119	19.9	5.55	99	20.6	5.82
72-74	52	18.4	6.12	40	22.2	5.21
75-77	2	5.0	1.00	3	24.7	1.25
78-80	1	21.0	--	2	8.5	4.50

CHAPTER FIVE: FALLIBILITY OF DEPENDENT AND INDEPENDENT MEASURES

In the weighted-center-level effects analysis, a typical regression equation takes the matrix form

$$Y = XB + e \, , \qquad (5\text{-}1)$$

where Y is n x 1, X is n x p, B is p x 1, e is n x 1, n is the sample size, and p the number of independent variables. We assume that the variables (X,Y) can be measured only with error. This section presents the approach in the NDCS to the problems engendered by fallibility of measures.

Many authors (e.g., Lord, 1960; Porter, 1967; Campbell and Erlebacher, 1970;, Cronbach et al., 1977) have written about effects of fallibility in independent variables on the ANCOVA strategies used to reduce the bias that results from the nonequivalence of treatment and comparison groups. In the present application, the dependent measures do not correlate highly with covariables at the child level. For the most part, statistical adjustment strategies were not necessary in the cognitive effects analysis. The one covariable used--race of child at the PPVT analyses--resulted only in very small adjustments and can be disregarded for all practical purposes. Therefore, the problems of bias usually encountered simply did not occur. However, there were substantial effects on estimating the magnitude and significance of center level main effects. This chapter focuses on these issues.

In order to deal with them, it is necessary to estimate the generalizabilities of major variables. The NDCS used variance components analyses to compute generalizabilities for nearly all measures (see Singer, Affholter, Goodrich, 1978). Unfortunately not all variables were analyzable in this way since specially designed "G-Studies" (see Cronbach et al., 1972) had not been performed. It was therefore necessary to make the estimates from the data as collected for the effects analyses, i.e., the data were neither crossed, nested or balanced. Nevertheless, most of the important generalizabilities were obtained via a variety of statistical methods. The most important results of these analyses are presented in Table 5-1.

First we consider the problems engendered by fallibility of the *dependent* measures. Errors in the dependent measures do not result in biased estimates of regression coefficients or the absolute amount of variance-explained. However, the total variance of the dependent measure is biased upward by the factor $1/\rho$, where ρ is the generalizability. Thus, the *fraction* of variance explained by the independent variables is reduced. The R^2 obtained in a regression can be mujltiplied by $1/\rho$ (approximatley 1.7 for the PSI and PPVT) to obtain better estimates of the fraction of *true* center-level variance explained by the independent variables. For large sample sizes, the effective ceiling for R^2 equals ρ, or approximately 60%, since even a perfect model cannot explain error in the independent measures. Because fallibility of the dependent measure attenuates R^2 statistical power to detect effects is reduced substantially. There is, unfortunately, little that can be done to adjust for this loss except to be aware of it in interpreting "null" effects. Many of these would not be null if the dependent variable had been measured more accurately.

On the other hand, it is easy to adjust regression findings for attenuation due to fallibility of *independent* variables. Suppose that the model is given by the matrix equation (5-1) and by the measurement equation

$$Z = X + U \qquad (5-2)$$

where e and U are uncorrelated errors, independent of X, and the covariance matrix of U is given by

$$\frac{1}{n} \; E\,(U'U) = \Sigma_u = \text{diag}\,(\sigma_1^2,\ \sigma_2^2,\ \sigma_p^2) \qquad (5\text{-}3)$$

Then the statistic

$$\hat{B} = P^{-1}\,(Z'Z)^{-1}Z'y \qquad (5\text{-}4)$$

is a consistent estimator of B,

where

$$P = I - N\,D\Lambda D\,(Z'Z)^{-1} \qquad (5\text{-}5)$$

and

$$\Lambda = I - \text{diag}\,(\rho_1,\ \rho_2,\ \ldots,\ \rho_p) \qquad (5\text{-}6)$$

is obtained from the generalizabilities $\{\rho;\ 1 \leq i \leq p\}$ of the components of X. A proof is provided by Fuller and Hidirglou (1978). This formula is simpler in the special case where components of Z are uncorrelated. Then

$$P = \text{diag}\,(\rho_1,\ \rho_2,\ \ldots,\ \rho_p) \qquad (5\text{-}7)$$

so that disattenuation of B is accomplished by applying the Lord-Porter disattenuation separately to each component. In the more general case the formula (5-4) must be used.

Consider the case where there are just two independent variables in the vector X. In this case we can make an explicit calculation. Suppose that the sample covariance matrix is given by

$$\frac{1}{N} Z'Z = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \qquad (5\text{-}8)$$

Then we obtain the following for the matrix P.

$$P = \begin{bmatrix} r^2 + \rho_1 & r(1-\rho_1) \\ r(1-\rho_2) & r^2 + \rho_2 \end{bmatrix} \qquad (5\text{-}9)$$

This correction is then used directly in Equation (5-4) to achieve disattenuated estimates for the regression coefficient vector B. This Technique can of course be generalized to the case of more than 2 independent variables.

We note, however, that we can disregard terms in r when r is small, so that in this case (5-5) can be replaced by the approximate formula (5-7). Examination of center-level correlations in Chapter 6 shows that for the models considered in the effects analysis, this assumption is tenable. Practically speaking, therefore, regression coefficients can be disattenuated one-by-one, dividing by the corresponding reliability.

The fallibility of independent measures also attenuates the absolute and relative amount of variance explained. The disattenuated variance-explained is consistently estimated by the form

$$(Y'Z)P^{-1}(Z'Z)^{-1}(Y'Z)', \text{ replacing} \qquad (5\text{-}10)$$

the standard, uncorrected version

$$(Y'Z)\ (Z'Z)^{-1}\ (Y'Z)' \qquad (5\text{-}11)$$

Because of the attenuation of variance explained, the power of statistical hypothesis tests for significance of effects is seriously diminished in the case of the staff-background variables (education, experience), where ρ ranges from .34 to .61. Classroom structure variables are, however, little affected.

Summary

Generalizabilities of major dependent and independent variables were computed. The generalizability of each of the major dependent variables was about .6, while that of independent variables ranged from .34 to .89. A major effect of the fallibility of these measures is to attenuate the fraction of variance explained and thus to reduce statistical power significantly. A formula was derived to achieve a matrix disattenuation of effect sizes for the independent variables. However, for practical purposes, variable-by-variable (scalar) disattenuation by the Lord-Porter formula was shown to be appropriate. The results of these analyses were not applied quantitatively in the actual analysis of effects. They were used, rather, to provide a background for the interpretation of effects analyses, presented in the next chapter.

CHAPTER SIX: CENTER LEVEL COGNITIVE EFFECTS ANALYSIS AND VALIDATION OF FINDINGS

Description of Variables

The variables used to analyze children's acquisition of cognitive skills in day care may be divided into four broad categories: dependent variables, classroom structure variables, caregiver qualifications variables and covariables. This classification is illustrated in Table 6.1.

Table 6.1

VARIABLES USED IN COGNITIVE ANALYSES

Dependent Variables	Classroom Structure Variables	Caregiver Qualifications Variables	Covariables
GPSI	NCHILD	EDUCATION	MOTHER'S ED
GPPVT	NSTAFF	HIGHEST DEGREE	FRACTION WHITE
MGPSI	RATIO	SPECIALIZATION	INCOME
MGPPVT	LCHILD	PREV DC EXP.	POVERTY INDEX
MDGPSI	LSTAFF	CENTER EXP.	NADULTS
MDGPPVT	LRATIO		PSIINT
RPPVT			PPVTINT

Dependent Variables

The dependent variables analyzed are generalized gain scores computed for the Preschool Inventory (PSI) and for the Peabody Picture Vocaublary Test (PPVT).* The construction and use of these gain scores are discussed earlier in this report. Generalized gain scores for the PSI and PPVT (GPSI and GPPVT) were computed for each child in the study for whom data were available for Fall 1976 and

---

*For a description of the PSI and PPVT and their psychometric properties see Bache, 1980b.

Spring 1977. Center-level mean gain scores (MGPSI and MGPPVT) and median gain scores (MDGPSI and MDGPPVT) were then calculated for each of the 57 centers in the study. In addition, adjusted PPVT gain scores (RPPVT) were computed, correcting for two covariables found to be slightly correlated with PPVT gains at the child level (see Note 1 following this chapter).

Classroom Structure Variables

The three classroom structure variables are group size (NCHILD), staff count (NSTAFF) and staff/child ratio (RATIO).* Observations were made of the number of children and staff present for each class during each hour throughout the day (9:00 AM to 4:00 PM), on each of two days in each of five months in the Phase III year. The classroom structure variables were computed by averaging measurements over both days, for each of the five time points in the day care year, for the three morning hours (because structured, teacher-directed activities are largely morning phenomena). In addition to these raw figures, transformed versions of these three variables were computed: $\log_{10}$ [group size] (LCHILD), $\log_{10}$ [number of staff] (LSTAFF), and $\log_{10}$ [staff/child ratio] (LRATIO). These measures were constructed by taking logarithms at the observation level and then aggregating the logged observations in the same manner as the unlogged observations (see Note 2).

---

*The classroom structure variables are described in more detail in J. Singer, D. Affholter, and R Goodrich, "Variance Components and the Dependability of Measures used in the National Day Care Study," Abt Associates, 1978.

Table 5-1

Estimated Generalizabilities of Major Variables

Measure	Description	Generalizability	Source
GPSI	PSI Generalized Change Score	.63	Singer, Affholter, Goodrich (1978) Ibid.
GPPVT	PPVT Generalized Change Score	.58	
LRATIO	Observed Staff-Child Ratio	.88	
LKIDS	Log Number of Children	.83	Ibid.
LSTAFF	Log Number of Staff	.89	Ibid.
PREV. DC. EXP.	Previous Day Care Experience of Staff	.34	Unpublished calculation
SPECIALIZATION	Staff Specialization	.61	Unpublished calculation
EDUCATION	Staff Education	.50	Crude estimate

Caregiver Qualifications Variables

Variables in this category relate to the education, training and experience of caregivers, computed at the center level. EDUCATION is measured as the mean years of education of staff observed in the morning hours, HIGHEST DEGREE is the mean highest degree (coded as an ordinal scale) of staff observed in the center. SPECIALIZATION is the proportion of staff present in the morning hours who have had specialization in subject matter related to children and child care. PREV DC EXP. reflects the mean years of experience observed caregivers have had in day care centers other than the center in which they presently work; CENTER EXP. is the mean number of years that caregivers present in the morning hours had worked in their current center.

Covariables

These variables were used in regressions to control for the effects of possible confounding factors, such as socioeconomic status (SES) and race. The intent here was not to estimate the effects of these variables, but rather to adjust for them in some manner if they were found to be associated with the gain scores. MOTHER'S ED, FRACTION WHITE, INCOME, NADULTS, PSIINT and PPVTINT were aggregated to the center level only over those children with either a valid GPSI or a valid GPPVT score. MOTHER'S ED is the mean years of education of the children's mothers; FRACTION WHITE is the proportion of white children. INCOME is the average income of families. POVERTY INDEX is the fraction of families in the center's 1970 Census tract with 1969 income below the poverty line. NADULTS is the average number of adults in the children's homes. PSIINT (PPVTINT) is the average number of days between the two testings upon which GPSI (GPPVT) is based. Although there were many more covariables that were available, these seven had the highest correlation with children's gain scores, and hence possibly required further examination. Recall (see Chapter Two) that at the child level, covariables were nearly completely uncorrelated with generalized change measures.

Analytic Issues

Before delving into the details of analysis, there are three key analytic concerns which must be addressed. First, there is the question of the unit of analysis, i.e., of the level of aggregation. Next, the sample must be defined; that is, it must be decided whether all 57 centers should be included in a single analysis, or whether the 49-center quasi-experiment and the eight-center Atlanta Public Schools substudy should be treated separately. Finally, a discussion of the approach to regression model construction may prove informative.

Unit of Analysis

Effects analyses were conducted entirely with variables aggregated to the center level, as discussed previously in Chapter Three. The center-level aggregates are based upon approximately 700 children for whom test data were available for the two time points and the staff members who were observed in the centers. Recall that the analyses were performed at the center level rather than at the classroom level for three reasons: (a) over the day care year, children within a given center switched classrooms, making it difficult to create classroom-level, child-oriented variables; (b) even if the classrooms were stable entities, certain classes have so few children that the aggregated measures would not be very stable; (c) the absence of significant effects of within-center (child-level) variables means that center aggregation effects are not problematic.

Sample Definition

A second issue is whether to analyze data from all 57 Phase III centers together, or whether to treat the 49-center quasi-experiment and the eight-center Atlanta Public School (APS) substudy separately. Both strategies have merits; for the NDCS cognitive main effects analysis, the former approach was taken.

The primary justification for analyzing the two groups separately is that, in addition to the experimental manipulation of staff/child ratio in all NDCS centers, in the APS substudy caregiver roles were also experimentally altered. Interest here, however, does not focus upon caregiver roles, but rather upon measures of classroom structure and caregiver qualifications; these variables should not have been influenced by the APS manipulation. It seems reasonable, therefore, to combine both groups in a single analysis.

In addition, the APS centers contribute valuable information to the overall description of day care. To analyze the two groups separately would alter the profile of NDCS centers (that is, change the variability patterns), which in turn could diminish the ability to detect effects. Moreover, since the data set is so small, even including all 57 centers, it seems wise to retain as much data as possible in a single analysis.

To ensure that this approach does not lead to misleading results, however, two checks have been inserted into the analysis. First, APS data have been separately analyzed (at the classroom level), in the detailed investigation of the APS.* Second, the data from the 49-center quasi-experiment were examined to see if the major effects found remained in the absence of the APS data. These validations, coupled with the major analysis of the 57-center data set, produced results which are stronger than the results of one of the individual examinations alone.

Approach to Regression Model Construction

Two primarly schools of thought on the topic of regression model construction are currently active. One enters all independent variables of interest into a model and then estimates effects on the basis of this single regression; the second takes a more parsimonious approach to model construction and selects independent variables paying careful attention to problems of multicollinearity. Each approach has its merits; the latter was chosen for the cognitive main effects analysis.

This approach has several advantages. Most important, when the independent variables are not truly independent (i.e., when they are multicollinear), the former approach may yield results that are uninterpretable, whereas a more exploratory and parsimonious strategy often leads to a series of regression equations which in combination indicate the relationships among independent and dependent variables. Many of the variables in the cognitive main effects analysis are collinear; therefore, this approach seems most reasonable.

---

*Goodrich, N., "The Atlanta Public Schools Day Care Experiment." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

An additional justification for this "lean" approach to model construction can be mounted on the basis of the ability to detect stable effects. By studying the changes in regression coefficient estimates as well as changes in estimates of standard errors, it is possible to acquire valuable information about the variability of the data. Moreover, this type of examination can aid in the detection of "outliers". For example, it is possible that the residuals from a regression on five independent variables could approximate a normal distribution, whereas the residuals from a regression involving only two of these regressors could highlight certain cases which appear to be atypical. The more inclusive strategy could mask possible outliers, while the more parsimonious approach leads to many more opportunities to detect effects of this sort.

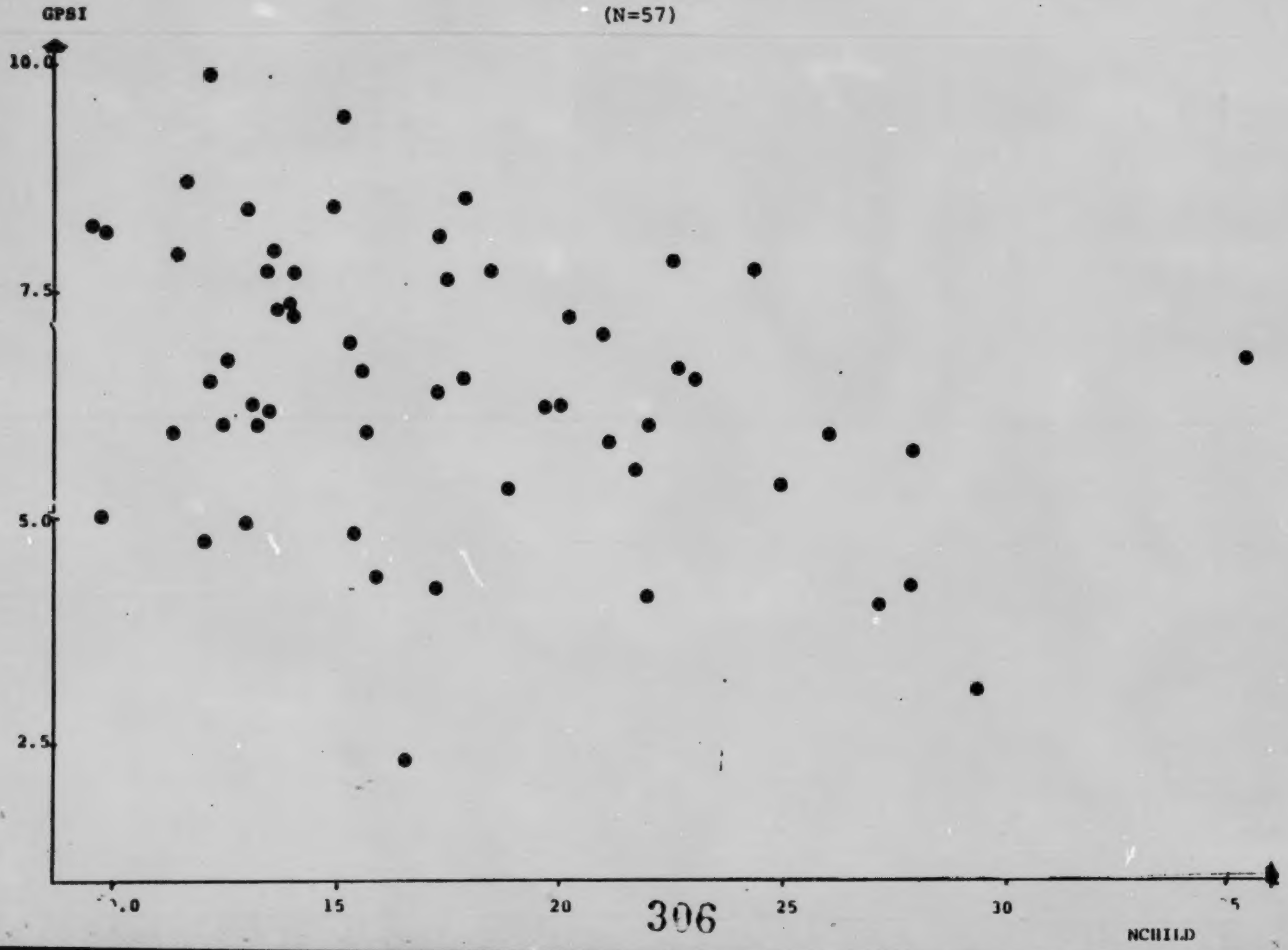
Data Analysis

Preschool Inventory (PSI)

The first stage in the analysis of the PSI generalized gain scores was to examine the two-way plots of MGPSI versus each of the available regressors. These graphs suggested that group size bore the strongest relationship to cognitive gains (see Figure 6.1). Staff count and staff qualifications also appeared to be associated with the gain scores, although the strength of these relationships was not as great. Moreover, the covariables seemed to be unrelated to MGPSI. This initial indication of a lack of covariable relationships to the gain scores was quite heartening. It suggested that, in addition to the lack of child-level covariable confounding, these data were not plagued with the confounding of such variables as racial composition and socioeconomic status with cognitive gain scores as so many other educational studies have been in the past.

Figure 6.1

GPSI VERSUS NCHILD
CENTER LEVEL
(N=57)



The correlation matrix among the variables (Table 6.2) not only confirmed what the scattergrams had suggested, but also served to highlight the interrelationships among the variables. That is, it illustrated the reasonableness and importance of conceptualizing the regressors as three clusters of independent variables (classroom structure, caregiver qualifications, covariables) rather than as 14 distinct variables. This formulation is based upon the relatively high degree of correlation within each variable set and the relatively low degree of correlation across the sets. Within the classroom structure variables, the correlation between group size and staff count is so high as to prohibit entering them both in a single regression without introducing severe multicollinearity problems. Centers in the NDCS were chosen to minimize the correlation between group size and ratio, making it possible to separate out the individual effects of each variable. Due to the correlations among these three variables, it would not be wise to include both group size and staff count, or both ratio and staff count, in a single model. However, since group size and ratio are relatively independent, it is reasonable to include both in a single regression.

Examination of the correlations among the staff qualifications variables also suggests problems of multicollinearity were several of them to be entered in one regression. The covariable cluster is also plagued by high correlations. This is due in part to the fact that five of the variables are SES-related. An attempt was made to create one SES variable through principle components analysis; this new variable did not correlate more highly with the gain score, however, than any of the individual covariables, lending further credence to the supposition that a covariable model was unnecessary.

Table 6.2

CENTER-LEVEL CORRELATION MATRIX[a]
(N=57)

	DEPENDENT MEASURES			CLASSROOM STRUCTURE			CAREGIVER QUALIFICATIONS					COVARIABLES					
	MGPSI	MGPPVT	RPPVT	LCHILD	LSTAFF	LRATIO	EDUC	HIGHEST DEGREE	SPEC-CIALI-ZATION	PREV. D.C. EXP.	CENTER EXPER.	MOM EDUC.	FRAC. WHITE	INCOME	POVERTY INDEX	NO. NADULTS	PSI INTERVAL
MGPSI	1	.33	.32														
LCHILD	-.33	-.30	-.27	1													
LSTAFF		-.31	-.26	.66	1												
LRATIO				-.26	.56	1											
EDUCATION		.31	.27				1										
HIGHEST DEGREE		.39	.36				.82	1									
SPECIALIZATION	.26						.34	.36	1								
PREVIOUS DAY CARE EXPERIENCE	.30					.21			.25	1							
CENTER EXPERIENCE							-.27		.26		1						
MOM EDUCATION		.27				-.22		.24			-.21	1					
FRACWHITE		.35					.24	.24			-.26	.40	1				
INCOME					-.28	-.31		.22			-.27	.79	.49	1			
POVERTY INDEX					.25	.32			.31		.30	-.56	-.48	-.67	1		
NUMBER NADULTS			-.21		-.28	-.31					-.32	.54		.61	-.49	1	
PSI INTERVAL						-.21	.31	.33				.22	.22		-.41		1
PPVT INTERVAL	.23	.21					.26	.31				.22	.24		-.36		.99

[a]Only correlations significant at the p<.15 appear on the table.

In addition to highlighting the correlations within the three sets of regressors, Table 6.2 also shows the correlations between the regressor sets. It appears that correlations among the sets of variables are relatively low, thus enabling different types of variables to enter into a single regression without introducing severe multicollinearity.

To get an initial picture of how the independent variables were related to the PSI gain scores, an ordinary least squares (OLS) stepwise regression was done. Of the 14 variables, four were entered into the equation (in order of entry): LCHILD, SPECIALIZATION, CENTER EXP, and PREV DC EXP. This regression reinforced early indications that both classroom structure and caregiver qualifications were associated with PSI gain scores. Note, however, that the staff/child ratio did not enter into this regression, suggesting that it is group size (and the corresponding staff count) that is associated with the gain scores, and not the ratio. A plot of the residuals versus the fitted values provided no evidence of heteroscedasticity. Moreover, this regression further strengthened the belief that it was unnecessary to include covariables in subsequent models, because they remained nonsignificant even when the effects due to classroom structure and caregiver qualifications were partialled out. Rather than dispense with the covariables altogether with the risk of arriving at misleading results, however, the covariables were simply set aside, while more exploratory types of analyses were conducted using the other two regressor sets. The covariables were reintroduced into several subsequent models to insure the "robustness" of the results to possibly confounding covariables.

Using biweighted regression,* an examination was made of each of the three logged classroom structure variables alone and in conjunction with various combinations of the caregiver qualifications variables. This analysis proved very enlightening for several reasons. First, it brought to our attention several data points which appeared to be atypical. Three centers (one in each site) consistently had biweighted weights in the range of 0.00 to 0.50; that is, they had large residuals in both the OLS and biweighted fits. On the basis of subsequent investigation, these three centers were set aside from the analysis of outliers. Second, this examination clearly illustrated the problems that multicollinear regressors tend to create. Third, it stressed the need not only to biweight analyses, but also to weight by the number of children in each center with valid gain scores.

Biweighted regression analysis proved to be a very useful analytic tool. For example, the results of biweighted regressions helped to illustrate the multicollinearity problems present in these data. Table 6.3 presents the OLS and biweighted regression coefficients for the consecutive stages of the stepwise regression previously described. The OLS coefficients and their respective t-statistics for the variable LCHILD appear to be relatively stable, as do the significance levels attached to them. Examination of the biweighted coefficients, however, suggests that the estimates are only stable when no more than one of the caregiver qualifications variables is present; that is, when the regressors are not highly correlated. In the presence of two qualifications variables the coefficient for LCHILD begins to fluctuate, and when the three collinear carriers are entered, the biweighted estimate is quite different from the OLS estimate. The t-statistics also fluctuate as a function of the multicollinearity: PREV DC EXP. is significant at first; it loses ground when SPECIALIZATION is included; then, in the presence of all three qualifications variables, PREV DC EXP. is significant again. This is due to the effect that multicollinearity has, not only on the estimated coefficients, but on their estimated variances as well. Therefore, due to the relatively high correlations among the qualifications variables, simpler models--suffering less from multicollinearity--were sought after.

---

*See Appendix 2 for a detailed description of biweighted regression.

Table 6.3

REGRESSION COEFFICIENTS FOR REGRESSION OF
MGPSI ON SELECTED VARIABLES

Ordinary Least Squares (OLS) and Biweighted Least Squares (BLS)
(N=57)

Carriers	OLS Coefficient	Standard Error	t	P Value	BLS Coefficient	R^2
LCHILD	-3.74	1.41	-2.66	.0099	-3.67	.11
LCHILD	-3.82	1.36	-2.82	.0067	-3.58	
PREV DC EXP.	0.16	0.07	2.30	.024	0.15	.19
LCHILD	-3.89	1.32	-2.95	.0049	-3.03	
PREV. DC EXP.	0.12	0.07	1.74	.084	0.12	
SPECIALIZATION	1.22	0.59	2.08	.040	1.28	.25
LCHILD	-4.16	1.29	-3.06	.0038	-2.44	
PREV DC EXP.	0.18	0.07	2.47	.016	0.18	
SPECIALIZATION	1.96	0.59	3.17	.0029	2.11	
CENTER EXP.	-0.17	0.12	-1.33	.190	-0.23	.31

Biweighting also reinforced the need for performing weighted least squares. Since the center-level aggregations were computed with varying numbers of children, the information contained in each case had a varying level of accuracy or stability. Furthermore, weighting by the appropriate number of children alleviates problems created by undue influence of small centers with few children, whereas equal weighting would tend to disproportionately favor those centers.

In summary, due to multicollinearity problems introduced when several qualifications variables were entered in the same regression, an effort was made to construct simpler models that would adequately portray the combined associations of classroom structure and staff qualifications with gain scores. In addition, three centers were identified as atypical (outliers), and set aside from subsequent analyses. Both weighted least squares and weighted-biweighted least squares techniques were used for the remaining analyses.

The next stage in analysis was to fit relatively simple models with each of the three classroom structure variables alone and in conjunction with PREV DC EXP. and SPECIALIZATION (Note 3). In addition, the pair of variables LCHILD and LRATIO were jointly examined in this context. The results of one set of these analyses (for LCHILD) appears in Table 6.4. Not only is group size negatively associated with cognitive gains (i.e., larger groups are associated with lower mean gains), but holding group size constant, higher gains are also associated with staff who have had previous day care experience and/or special courses in child development. The stability of the biweighted coefficients reinforces the strength of this finding.

Results for the other two classroom structure variables are nonsignificant; however, the signs are in the expected direction (Note 4). Furthermore, when both group size and ratio are entered together, group size remains significant while ratio never approaches significance. When ratio alone is entered, with or without the two staff

Table 6.4

RESULTS OF WEIGHTED AND WEIGHTED-BIWEIGHTED REGRESSIONS ON MGPSI FOR SELECTED VARIABLES

Independent Variables	Weighted Regression Coefficient	Standard Error	t	P Value	Weighted Biweighted Regression Coefficient	R^2 (for weighted regression)
LCHILD	-3.79	1.38	-2.74	.0083	-3.40	.13
LCHILD	-3.81	1.34	-2.84	.0065	-3.38	
PREV DC EXP.	.16	.08	2.02	.046	.15	.19
LCHILD	-4.31	1.33	-3.24	.0025	-3.13	
SPECIALIZATION	1.35	.53	2.55	.013	1.57	.23

qualifications variables, not only is its contribution nonsignificant, but PREV DC EXP. and SPECIALIZATION also are nonsignificant. This suggests that group size and caregiver characteristics ought to be considered jointly.

To check for the stability of these results in the presence of covariables, weighted and weighted-biweighted regressions were estimated using the three sets of regressors in Table 6.4, in conjunction with each of the covariables. In every instance, the covariables were nonsignificant; furthermore, they hardly exerted any influence on the estimates of the other coefficients. This result strengthened the earlier belief that the inclusion of covariables was not necessary.

In order to protect these results from the possibility that mean center gain scores are unduly influenced by extreme individual scores within a given cetner, the above regressions were re-estimated using median PSI gain scores (MDGPSI). The results were somewhat weaker than those using the mean gains; however, they indicated the same findings as previously noted.

A potential threat to the validity of conclusions concerning cognitive gains of children in the NDCS sample can be argued from differential attrition in the sample, by center. If attrition is associated with some quality of "thrivingness" in a day care seting, if "thrivingness" is associated with cognitive gains, and if attrition is also associated with the policy variables under consideration, then any conclusions regarding the effects of policy variables on cognitive gains may be distorted by differential attrition by center. The direction of distortion depends upon the empirical relationships among attrition, cognitive gains, and policy variables; that is, purported effects may be either attenuated or exaggerated.

A simple example can be developed to illustrate both possibilities. Suppose that attrition and cognitive gains are associated, and that centers with higher attrition rates are composed disproportionately of "thrivers". Suppose further that attrition and group size are positively associated, that classes with larger enrollments tend to have higher attrition rates. (For instance, signs of "nonthrivingness" are more likely to be observed by parents of children in large classes, with resulting higher frequencies of removal of "nonthrivers" from the day care context.) If group size is negatively associated with cognitive gains, ceteris paribus, then that result is somewhat attenuated by attrition: disproportionate representation of "thrivers" in large classrooms would dampen negative effects of group size on cognitive gains, since "thrivers" are expected to have higher gain scores than "nonthrivers."

By changing one relationship in the hypothetical argument given above, the consequences of differential attrition change. Suppose, now, that group size is negatively associated with attrition (larger classes have lower attrition rates, or larger classes effectively mask signs of "nonthrivingness"). If the other relationships postulated in the paragraph above still hold, then the negative effects of group size, ceteris paribus, are somewhat exaggerated by attrition. Disproportionate representation of "nonthrivers" in larger classrooms exacerbates the negative effects of group size on gains since "nonthrivers" are expected to have lower gain scores.

Attrition rates (fraction of children in the NDCS sample attrited between Fall 1976 and Spring 1977) were calculated for each center and correlated with the major policy variables. The results appear in Table 6.5. Negative effects of group size may be somewhat exaggerated, and positive effects of caregiver characteristics may be somewhat attenuated by differential attrition--but none of the relationships are strong enough to attain statistical significance at the .10 level. Apparently, a serious threat to the validity of the NDCS cognitive effects conclusions cannot be mounted on the basis of differential attrition.

Another possible threat to the findings in this analysis is disproportionate influence of the few data points on fitted equations. Further examination of PREV DC EXP. indicated that the significant results for this variable were mainly due to four centers in the study. These centers had the highest values for PREV DC EXP. and correspondingly high PSI gain scores. Although the direction of relationships between cognitive gains and previous day care experience remains the same when these four centers are eliminated from the analysis, the strength of the association is substantially weakened and this variable is no longer significant. This behavior is illustrated in Figure 6.2. Examination of this graph of MGPSI versus PREV DC EXP. reveals a strong positive

Table 6.5

CORRELATIONS BETWEEN FRACTION ATTRITED AND OTHER VARIABLES

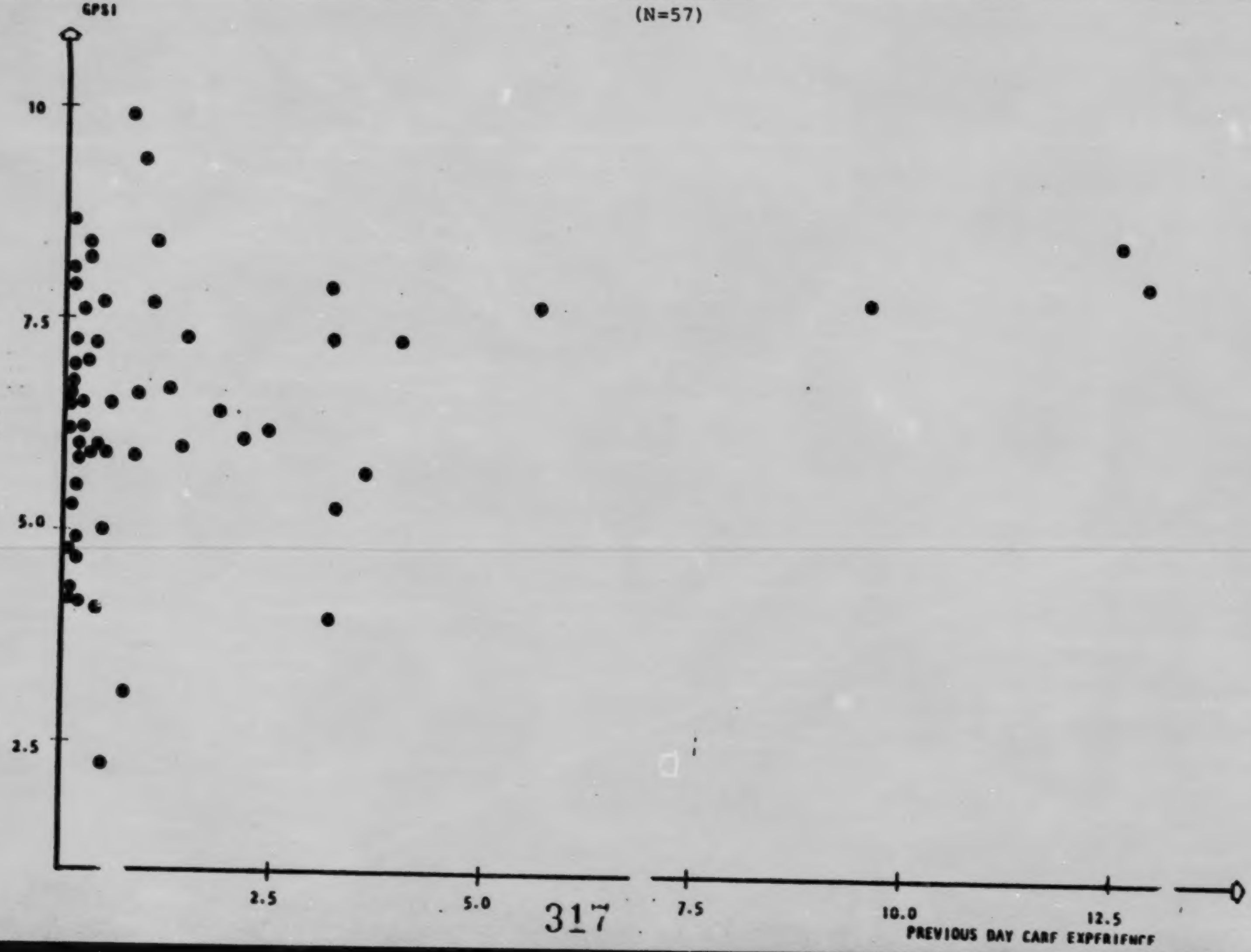
Center-Level Correlations
(N=57)

Variables	Correlation
MGPSI	.17
LCHILD	-.10
LSTAFF	-.18
LRATIO	-.12
HIGHEST DEGREE	-.08
EDUCATION	-.05
SPECIALIZATION	-.20
PREV DC EXP.	-.18

slope. If the centers with the extremely large values of PREV DC EXP. (greater than 5.00) are set aside, the effect is severely attenuated and is no longer significant. This is not to say that these centers are outliers which should be removed from the analysis, and that there is no association between previous day care experience and PSI gain scores. Data in the extremes of the independent variables are usually considered high leverage points; that is, they exert more influence upon the regression estimation procedure than do the points in the middle of the distribution. If there was simply one center with an extremely high value on this variable, it would probably have been considered an outlier and set aside. Since there are four centers with this characteristic, however, it is not desirable to call them all outliers, because we may be eliminating valuable data. The problem becomes one of not having a sufficient number of centers with large amounts of previous day care experience to draw solid conclusions. Therefore the findings for this variable cannot be considered definitive.

GPSI VERSUS PREVIOUS DAY CARE EXPERIENCE
CENTER LEVEL

(N=57)



As an additional check on the stability of the findings, a type of cross-validation was done. The sample was partitioned into four sections: APS, Atlanta Non-APS, Detroit and Seattle. Ideally, separate analyses would have been conducted within each subsample. However, since each subsample was too small to support separate analysis, the cross-validation involved analyzing different combinations of subsamples, omitting a different one (or two) in each case. A portion of the results of these validations appear in Table 6.6. On the whole, these results support the hypothesis that group size and specialization are associated with cognitive gains. There are refinements of this hypothesis, however, that need to be made. First, there are subsamples of the data in which ratio is a significant variable, most notably when Atlanta Non-APS and/or Seattle is removed. There is no immediately apparent reason for this behavior. Moreover, ratio is significant when APS is removed from the sample, but the coefficient is so unstable when biweighted that the actual significance may be doubtful. It is important to note that in those subsamples for which ratio is significant, group size alone is also significant. The essentially null findings with respect to ratio stands, therefore, through this validation.

The results for group size and caregiver qualifications are not as clearly defined. First, the group size effect appears to be the strongest in the APS/Detroit subsample, while the caregiver qualifications variables are weakest in this subsample. In the complementary subsample, Atlanta Non-APS/Seattle, exactly the reverse effects emerge (strong qualifications and weaker group size results). There is no a priori reason for this particular partition--APS and Detroit, Atlanta Non-APS and Seattle--and while it may produce a distortion of the effects, there is no apparent explanation for this behavior. When all Atlanta centers are

Table 6.6

REGRESSION COEFFICIENTS MGPSI ON SELECTED POLICY VARIABLES

(Regressions are weighted by number of children
in center with valid gain scores)
(N=54)

CODE: APS = APS
NAPS = Atlanta Non-APS
ALL = All Sites
DET = Detroit
SEA = Seattle

Sites	LCHILD	SPECIALIZATION	LRATIO	SPECIALIZATION
ALL [initial]	-4.31	1.35	1.98	1.04
[t]	(-3.24)	(2.55)	(1.19)	(1.84)
[biweighted]	-3.13	1.57	.335	1.87
NAPS,DET,SEA	-2.60	1.34	.508	1.11
(49 Centers)	(-2.08)	(2.49)	(.369)	(2.00)
	-.153	1.86	.091	1.78
APS,NAPS,DET	-4.84	1.48	1.85	1.13
	(-3.03)	(2.46)	(0.90)	(1.72)
	-3.26	1.73	0.12	2.07
APS,DET,SEA	-5.46	0.98	6.09	.301
	(-3.13)	(1.41)	(1.78)	(.359)
	-5.29	1.01	4.03	.939
APS,NAPS,SEA	-4.68	1.62	2.51	1.24
	(-2.83)	(2.28)	(1.44)	(1.67)
	-1.72	2.17	.066	5.82
APS,NAPS	-5.73	1.93	2.56	1.40
	(-2.62)	(2.15)	(1.09)	(1.46)
	-2.34	2.29	.435	2.22
DET,SEA	-2.81	0.38	.533	.101
	(-1.68)	(0.45)	(.188)	(.110)
	-2.75	0.42	.497	.246
NAPS,SEA	-2.81	1.46	1.35	1.67
	(-1.68)	(2.35)	(1.03)	(2.48)
	-2.75	2.17	.687	2.37
APS,DET	-6.62	1.12	11.06	-0.84
	(-2.72)	(1.26)	(1.67)	(-.070)
	-6.78	1.15	10.50	.174

removed, the caregiver specialization effect disappears. This is due, in part, to the fact that the centers in Atlanta have caregivers with more specialized training than in the other two sites. When these centers are removed from the sample, therefore, variability in qualifications is markedly reduced, decreasing their strength as regressors. Hence, there are certain partitions of the sample which will alter the distributions of the independent variables which, in turn, may increase or decrease their predictive power. Although the magnitude and relative significance of certain effects may vary from partition to partition, the direction of the effects is never reversed. The basic conclusions do not appear to be threatened.

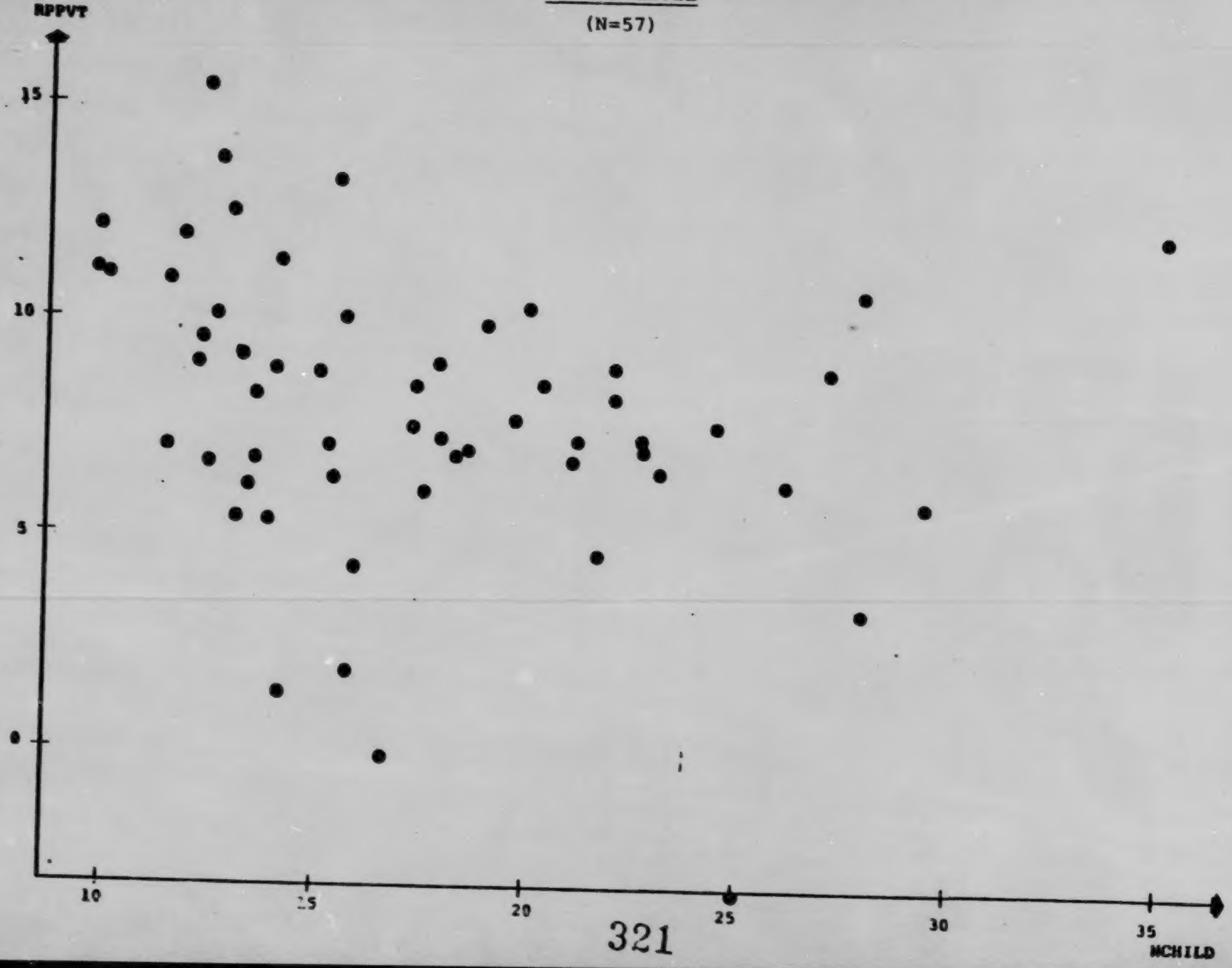
Peabody Picture Vocabulary Test (PPVT)

The analysis of PPVT gain scores was conducted in parallel to that of the PSI. Analyses used both the aggregated PPVT gain scores (MGPPVT) and the covariable-adjusted gain scores (RPPVT), as discussed above.

Preliminary examination of the PPVT gain scores through plots nad correlations suggested that as with the PSI, the variables GROUP SIZE and STAFF COUNT are strongly associated with gains. For example, a plot of RPPVT versus NCHILD appears in Figure 6.3. The center-level correlations between the PPVT gains and the independent variables appear in Table 6.2. Examination of this correlation matrix shows that not only are the classroom structure variables associated with gains, but the two caregiver education variables (HIGHEST DEGREE and EDUCATION) appear to be strongly related to PPVT gains. Furthermore, unlike the PSI, caregiver experience and specialization do not seem to be associated with PPVT gains. Although two of the covariables (FRACTION WHITE and NADULTS) are moderately correlated with MGPPVT,

RPPVT VERSUS NCHILD
CENTER LEVEL

(N=57)



the covariable-adjusted RPPVT does not suffer from this problem. All analyses have been conducted using both these gain scores; however, attention here is focused primarily upon the RPPVT.

Exploratory OLS and biweighted regressions were run using group size, staff count, and the caregiver education variables. Staff count performed slightly better than group size in these analyses, but because of the previously mentioned correlation between the two, the implications remain the same--classrooms with fewer children and staff tend to be associated with higher cognitive gain scores. The results for caregiver education variables are not as clearly defined. When the caregiver education variables were entered into regressions, the biweighted coefficients became very unstable. Examination of the regression residuals and weights from the biweighted estimation showed three centers with consistently large residuals and, hence, consistently low weights. These three centers were those previously determined to be outliers in the PSI analysis. This problem is illustrated in Figure 6.4. In this plot of RPPVT vs. EDUCATION, the centers which are circled (with arrows pointing towards them) are the three centers in question. On this graph there is a distinct positive slope; however, without these three points, the slope flattens out considerably. Hence, these three centers were again deleted from further analyses.

Weighted regression and weighted-biweighted regressions were then estimated using group size and staff count separately, and each in conjunction with the caregiver qualifications variables. Selected results of these analyses appear in Table 6.7. Note that the R^2s are substantially lower than those obtained for the PSI. In most instances, however, effects are in the same direction. The PPVT

Figure 6.4

RPPVT VERSUS EDUCATION
CENTER LEVEL

(N=57)

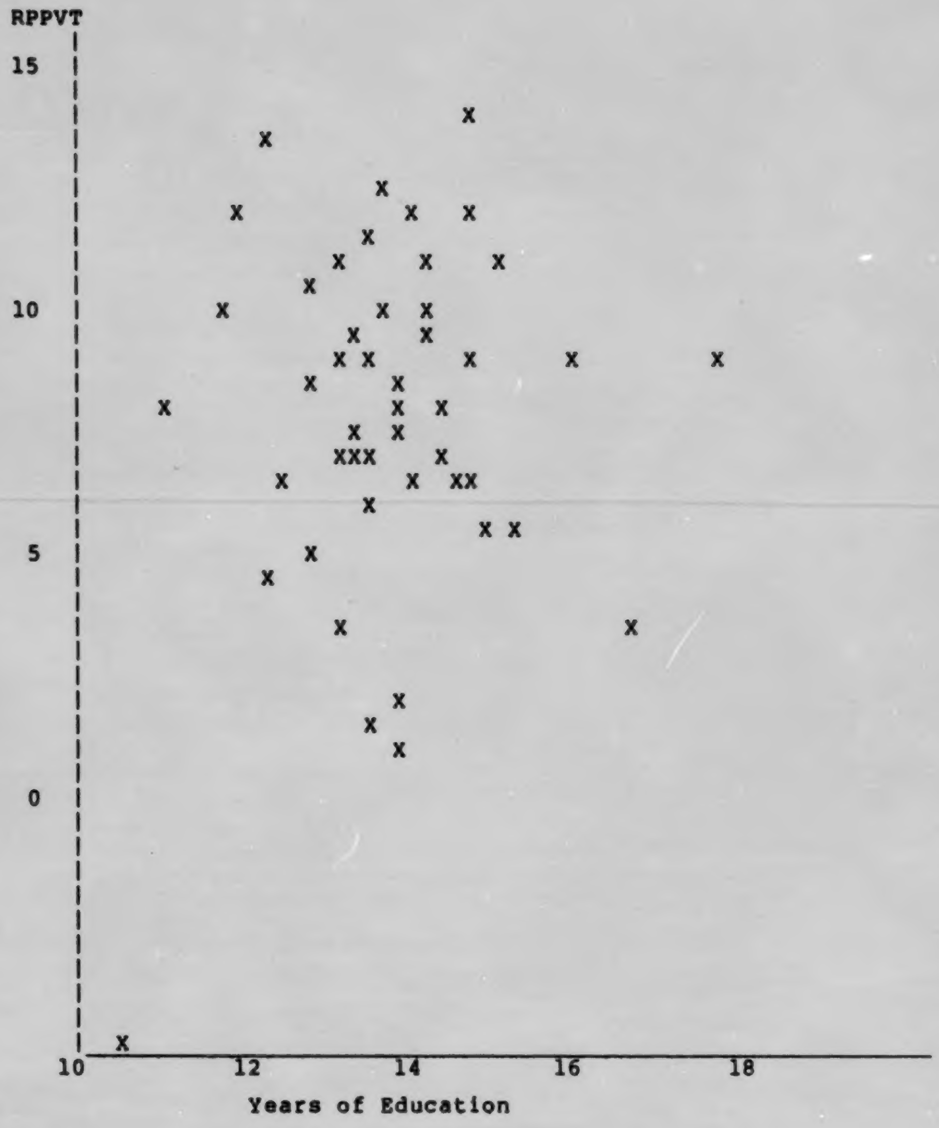


Table 6.7

RESULTS OF WEIGHTED AND WEIGHTED-BIWEIGHTED REGRESSIONS ON PPVT GAINS FOR SELECTED VARIABLES

(N=54)

Dependent Variable	Independent Variables	Weighted Regression Coefficient	Standard Error	t	P Value	Weighted Biweighted Regression Coefficient	R^2 (for Weighted Regression
MGPPVT	LCHILD	-5.20	2.80	-1.86	.065	-8.59	.06
	LSTAFF	-4.84	2.49	-1.94	.055	-5.72	.07
	LSTAFF	-5.54	2.50	-2.21	.030	-6.53	
	PREV DC EXP.	.24	0.16	1.52	.13	.29	.11
	LSTAFF	-4.83	2.55	-1.90	.060	-6.30	
	SPECIALIZATION	-.03	1.14	-.03	.97	.93	.07
	LSTAFF	-5.24	2.49	-2.10	.038	-6.56	
	EDUCATION	.45	.35	1.28	.20	.50	.10
	LSTAFF	-5.13	2.46	-2.08	.040	-6.36	
	HIGHEST DEGREE	1.33	.86	1.54	.13	1.40	.11
RPPVT	LCHILD	-4.09	2.64	-1.55	.12	-7.25	.04
	LSTAFF	-3.36	2.36	-1.42	.16	-3.38	.04
	LSTAFF	-4.07	2.37	-1.72	.088	-4.27	
	PREV DC EXP.	.25	.15	1.63	.11	.26	.08
	LSTAFF	-3.65	2.40	-1.52	.13	-4.08	
	SPECIALIZATION	.22	1.07	.66	.52	.28	.05
	LSTAFF	-3.56	2.39	-1.49	.14	-3.77	
	EDUCATION	.22	.34	.66	.52	.28	.05
	LSTAFF	-3.51	2.38	-1.48	.14	-3.71	
	HIGHEST DEGREE	.68	.83	.82	.58	.75	.05

results may be best viewed as confirming those obtained for the PSI rather than as solid evidence in their own right. Also note that the only variables that approach significance are GROUP SIZE, and STAFF COUNT. Although PREV DC EXP. has the highest t-statistic of any of the caregiver qualifications variables, this is primarily due to the same four centers that created a significant effect for this variable for the PSI. Furthermore, the variables SPECIALIZATION, EDUCATION and HIGHEST DEGREE do not appear to be significantly related to the PPVT gain scores. On the basis of these analyses, then, the size of the classroom grouping (as measured by group size and staff count) is the only factor that is associated with PPVT gains.

As was done for the PSI, an analysis of the median PPVT gains was performed. Although these results were somewhat weaker than those for mean gains, they remained the same.

Summary

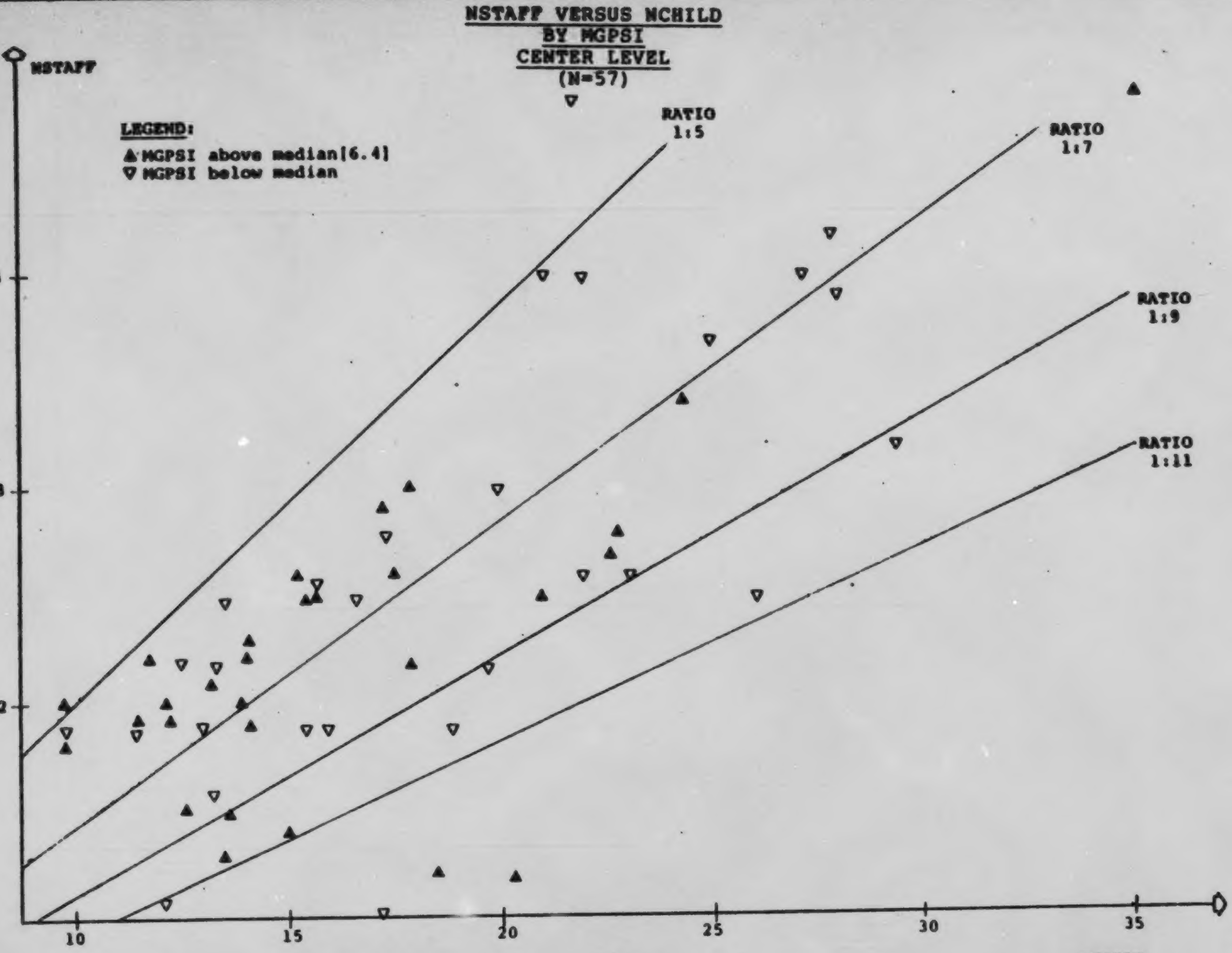
The data provide clear evidence that the structure of the day care classroom--the configuration of children and caregivers--is associated with children's acquisition of certain skills, as measured by the PSI and PPVT. Moreover, there is evidence that certain aspects of caregiver qualifications are likewise related to child outcomes.

The results of these analyses suggest certain configurations of caregivers and children to be encouraged, other considerations for policymakers not withstanding. Certers in which classes were composed of small groups of children interacting with few caregivers had higher cognitive gain scores on the average than centers where the classes were comprised of large numbers of children with correspondingly more caregivers. Within a range of reasonable ratios, it is the sheer magnitude of the classroom--the number of people present, not the prevailing staff/child ratio--that is most strongly associated with gains. To illustrate the association between classroom configurations and cognitive gains, a plot of group size and staff count (and hence ratio) with the cognitive gain scores was constructed (see Figures 6.5 and 6.6). The rays drawn on the graph are lines of constant ratio. To reduce a three-dimensional graph to two dimensions, the distributions of MGPSI and RPPVT were divided at the median and two different symbols were used to represent the third dimension.* As previously indicated, on the whole, centers comprised of classrooms with small groups of children and few caregivers appear to be better for children than those in which classes have more caregivers and correspondingly larger groups with similar staff/child ratios. The distribution of gain scores does not appear to be affected by slight changes in the staff/child ratio (within this narrow band of ratios). That is, there does not appear to be a concentration of centers with higher gains in the higher ratio centers. In addition to highlighting earlier results, the plots illustrate that it is virtually impossible to separate out the effects of group size and staff count; it is essential that they be considered jointly.

In addition to the conclusions about classroom groupings, analyses provided information aobut the associations between caregiver qualifications and cognitive gains. First, although the amount of previous day care experience was significantly related to cognitive gains, this result is

---

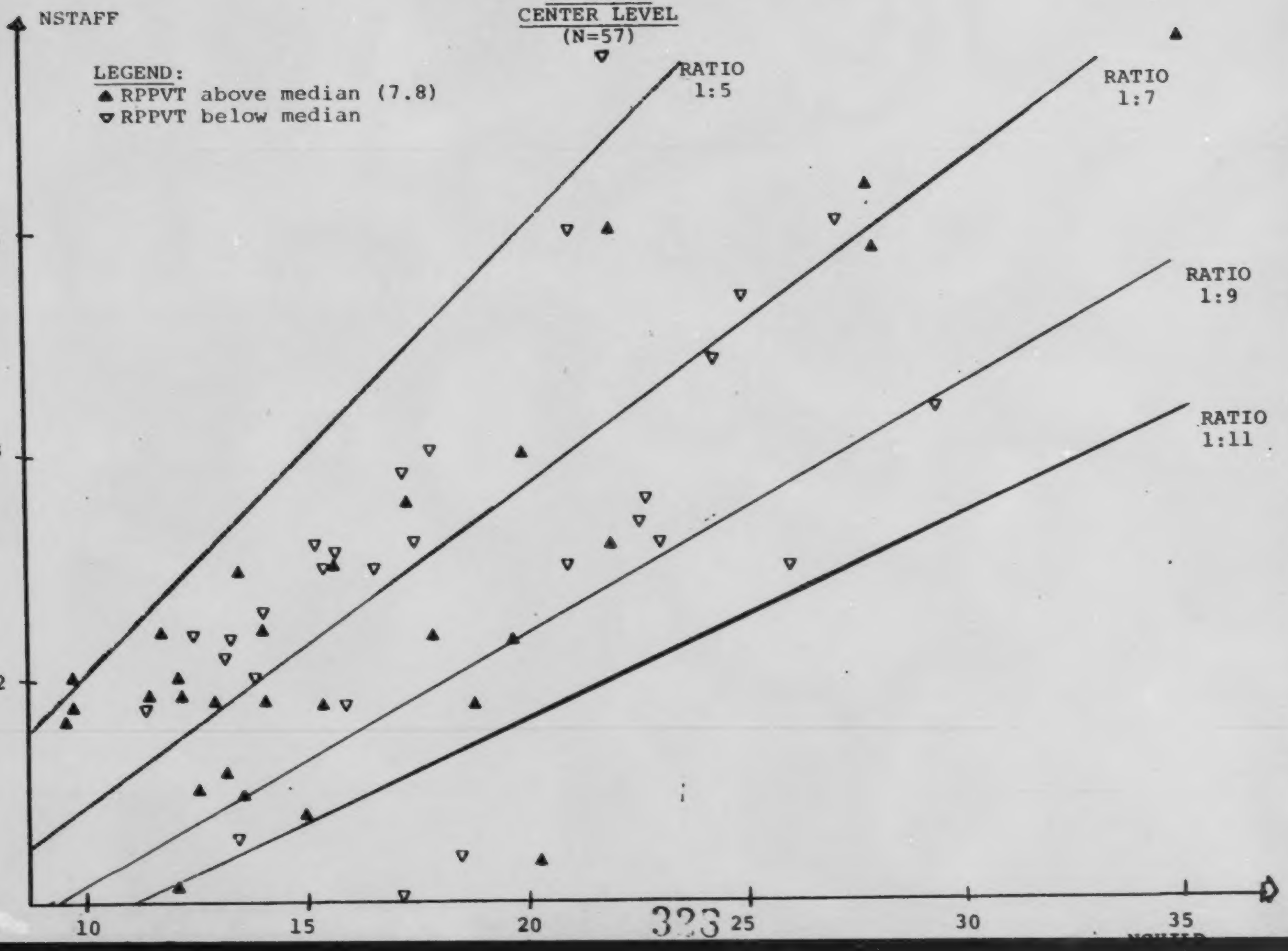
*A word of caution must be given here before interpreting these graphs. Because the gain scores have been recorded only as above or below the median, a great deal of information about their variability has been lost. While they show whether a center's gain score was above or below the median, they do not show how far above or below.



NSTAFF VERSUS NCHILD
BY MGPSI
CENTER LEVEL
(N=57)

Figure 6.6

NSTAFF VERSUS NCHILD
BY RPPVT
CENTER LEVEL
(N=57)



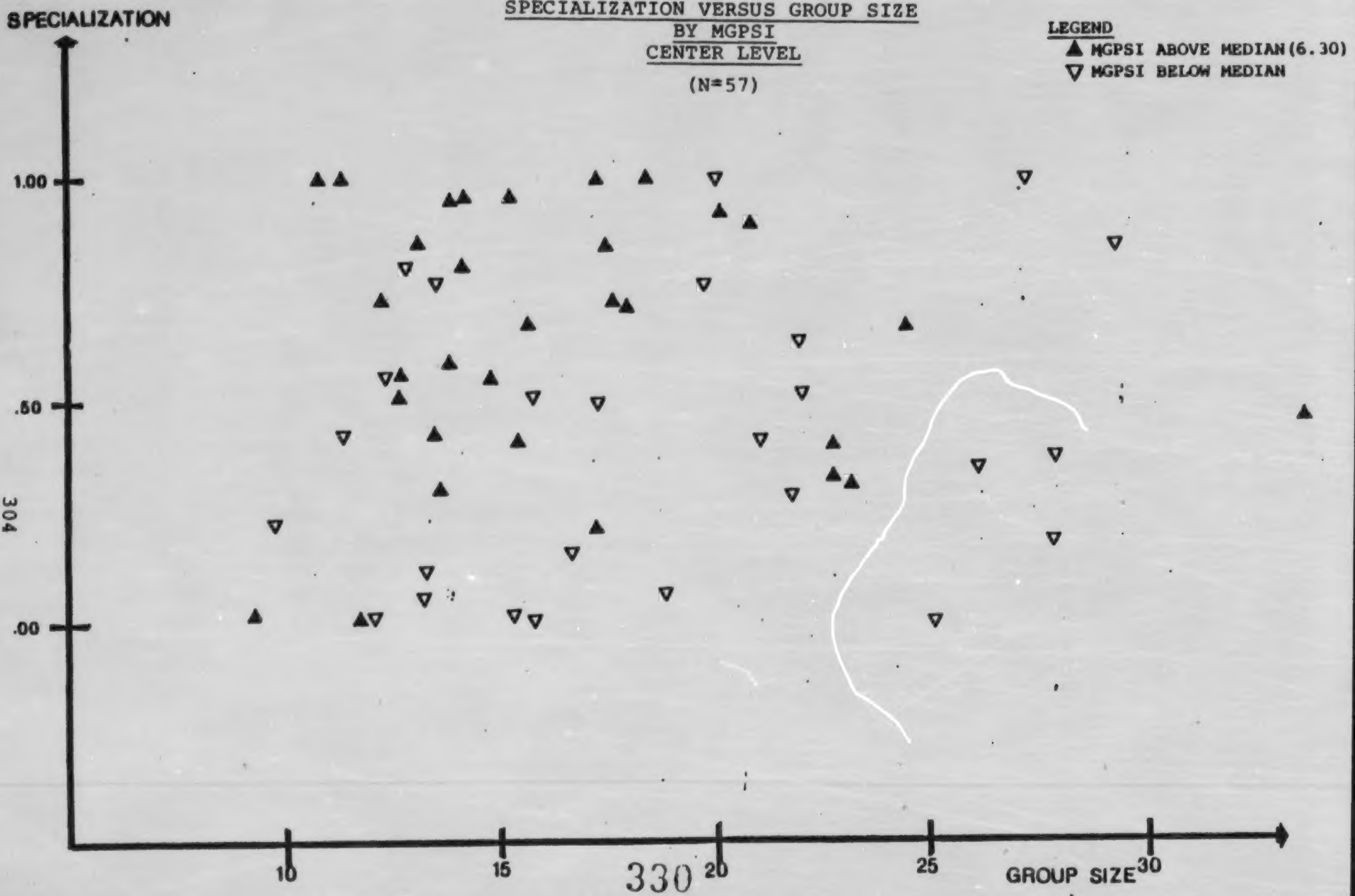
due primarily (as previously described) to four centers whose caregivers had high amounts of previous experience and whose children had high cognitive gain scores. On the basis of these analyses then, one cannot say very much about the effect of previous day care experience upon gain scores. Second, the amount of formal education by itself bore no significant relationship to cognitive gains. Third, caregiver specialization in child-related fields such as developmental psychology, early childhood education or special education was associated with children's PSI gain scores. Note, however, that the fraction of caregivers with specialized education in a center appeared to be unrelated to PPVT gains.

With the results of classroom groupings and caregiver specialization in mind, we now consider whether these are two independent dimensions of quality of care or if they must be considered jointly. A simple example will illustrate the difference between these two possibilities. It has been shown that classrooms comprised of smaller groups and caregivers with specialized education or training are both associated with higher gains. It could be the case that specialized caregivers would be effective with groups of varying sizes while caregivers without specialization would only be effective with smaller groups. On the other hand, it might be that all caregivers (regardless of their level of specialization) would be more effective with smaller groups and that specialized caregivers are more effective than non-specialized ones with groups of any size. In the former case, there is said to be a statistical interaction between specialization and group size; in the latter case, the effects are said to be additive. As shown in Figure 6.7, the effects of specialization and group size appear to be simply additive. In this graph of specialization, group size and PSI gain scores, there is a large concentration of high gain scores in those centers where

Figure 6.7

SPECIALIZATION VERSUS GROUP SIZE
BY MGPSI
CENTER LEVEL

(N=57)



there are smaller groups and more specialized caregivers. Furthermore, the benefits of small group size appear to hold regardless of the level of specialization and the benefits of specialization seem to remain regardless of the group size. It appears, then, that specialization and group size may be considered important and relatively independent dimensions in day care that influence children's acquisition of certain skills, as measured by the PSI and PPVT (Note 5).

NOTES TO CHAPTER 6

1. The RPPVT was computed as:

   RPPVT = GPPVT - 1.852 (FRACTION WHITE) - 1.05 (NADULTS)

   The correlation between RPPVT and GPPVT was approximately .97 and, in every case, analyses conducted upon each variable led to the same results.

2. This transformation was done for several reasons. First, when data consist of counts, a log transform often mitigates undesirable distributional properties in the sample data. Second, since ratio is the quotient of NSTAFF divided by NCHILD, ratio in the log metric is a linear combination; in fact, the simple difference of the other two logged counts. Third, the zero-order correlations among these three variables are slightly lower in the log scale than in the raw scale, thus helping to reduce the problem of multicollinearity. Because the correlation between the logged and unlogged variables is high (on the order of .95 to .99), there is little difference between results using the two sets of measures. For analytic purposes the logged variables are used; for displays, the untransformed data are shown to provide easy interpretation.

3. CENTER EXP was dropped from this analysis for two reasons. First, investigation indicated that slight errors had been made in the creation of the variable. These mistakes were exacerbated by improperly aggregating to the center level. Second, not only was this variable an insignificant predictor of MGPSI individually, but it was never significant when included in regressions with the classroom structure variables.

4. One possible explanation for the "superior" performance of group size over staff count is the relative lack of variation in staff count compared to variation in group size.

5. Although it is possible to state that centers with smaller classroom groupings and more specialized caregivers tend to have higher cognitive gain scores, it is difficult to state the precise numbers of children and caregivers in those small configurations, and the fraction of staff with specialized training, that should be encouraged. The variables analyzed were center-level averages of many classroom-level observations of group size, staff count and ratio, and, as such, they are continuous variables. The data were examined for nonlinear trends, or possible inflection points, and

none were found. That is, the above mentioned relationships appear to be linear, within the range of our data. Because of these effects, explicit "ideal" ranges for these variables cannot be given here.

6. Outliers may take many forms; they may be outliers in either the conditional distribution (that is, deviations from the regression surface), the unconditional distribution (extreme values of the observed values, relative to the rest of the sample), or both. Biweighted regression essentially attempts to deal with deviations from the regression surface that are large. Other "outliers" (and some analysts might debate the propriety of that label here) may "unduly" influence the regression fit without necessarily being extreme deviations from the fitted surface. Such observations are said to have high "leverage," and another set of diagnostic tools is required for examining data with potential problems from a few observations with high leverage (see Hoaglin and Welsch, 1978).

7. Since hypothesis tests are important in most research, the NDCS included, and since the sampling properties of biweighted regression estimates are not analytically tractable (Gross, 1977), researchers with these estimation problems face something of a dilemma. OLS techniques yield the required tests, but their propriety or accuracy may be questionable; biweighted estimates are more appropriate point estimates, but their sampling properties (and therefore their statistical significance) is unknown. This dilemma was attacked in the NDCS by exploratory analysis and model-building stages, checking results periodically and finally by comparing OLS-estimated coefficients with biweighted coefficients. Resources are rarely, if ever, sufficient for checking all OLS results step by step, and it is a moot point whether an observed biweighted-OLS difference is to be considered large or important. Nevertheless this technique proved to be very useful in the NDCS: erroneous data were identified and corrected, and outliers were discovered, scrutinized, and rejected from further statistical analyses.

REFERENCES

Affholter, D.P. and Bache, W.L., III, "The Application of Generalizability Theory to Observation Based Dependent and Independent Measures," Annual Meeting of A.E.R.A., San Francisco, 1979.

Anderson, T.W., An Introduction to Multivariate Statistical Analysis. New York: Wiley, 1958.

Andrews, D.F., Bickel, P.J., Jampel, F.R., Huber, P.J., Rogers, W.H., and Tukey, J.W., Robust Estimates of Location, Princeton, N.J.: Princeton University Press, 1972.

Arora, S.S., "Error Components Regression Models and Their Applications," Annals of Economics and Soc. Meas., 1973, 2, pp. 451-462.

Astrom, K.J., Introduction to Stochastic Control Theory. New York: Academic Press, 1970.

Bache, W.L. "Comparing Alternative Measures of Classroom Composition." National Day Care Study Measurements and Methods. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980a.

Bache, W.L., III, "A Psychometric Analysis of NDCS Phase III Child Test Battery," National Day Care Study Measurements and Methods. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980b.

Baram, Y., Non-Stationary Model Validation from Finite Data Records, Unpublished manuscript.

Box, G.E.P., and Jenkins, G.M., Time Series Analysis. San Francisco: Holden-Day, 1970.

Bryk, A.S., "An Investigation of the Effectiveness of Alternative Statistical Adjustment Strategies in the Analysis of Quasi-Experimental Growth Data." Doctoral Thesis. Graduate School of Education, Harvard, 1977.

Bryk, A.S., "Extensions of the Value Added Methodology," Mimeograph, Huron Institute, Cambridge, Mass., 1977.

3ryk, A.S., and Weisberg, H.I., "Use of Nonequivalent Central Group Design when Subjects are Growing," Psychological Bulletin, 1977.

Campbell, D.T., and Erlebacher, A., "How Regression Artifacts in Quasi-experimental Evaluations Can Mistakenly Make Compensatory Education Look Harmful," In J. Hellmuth (ed.), Compensatory Education: A National Debate--The Disadvantaged Child, Vol. III. New York: Brunner/Mazel, 1970.

Campbell, D.T., and Stanley, J.C., "Experimental and Quasi-Experimental Designs for Research," Handbook of Research on Teaching. AERA, 1963.

Cronbach, L.J., "Research on Classrooms and Schools: Formulation of Questions, Design and Analysis," Occasional Papers of the Stanford Evaluation Consortium, Stanford University, July 1976.

Cronbach, L.J., and Furby, L., "How We Should Measure Change--or Should We?" Psychological Bulletin, 1970, 74, 68-80.

Cronbach, L.J., Gleser, G.C., Nanda, H., and Rejartnam, N., The Dependability of Behavioral Measurements. New York: Wiley, 1972.

Cronbach, L.J., Rogosa, D.R., Floden, R.E., and Price, G.G., "Analysis of Covariance in Non-randomized Experiments: Factors Affecting Bias," Stanford Evaluation Consortium, 1977.

Duncan, O.D. and Davis, B. "An Alternative to Ecological Correlation," American Journal of Sociology, 1953, 64, pp. 610-625.

Dunn, L.M., "Peabody Picture Vocabulary Test," Circle Pines, Minn.: American Guidance Service, 1959.

Educational Testing Service, Preschool Inventory Revised Edition--1970: Handbook. Princeton: Educational Testing Service, 1970.

Fuller, W.A., and Hidiroglou, M.A., "Regression Estimation after Correcting for Attenuation," J.A.S.A., 1978, 73, 99-112.

Gehlke, C. and Biehl, R., "Certain Effects of Grouping Upon the Size of the Correlation Coefficient in Census Tract Material," J.A.S.A., 1934, 29, pp. 169-170.

Goldberger, A.S., "Selection Bias in Evaluating Treatment Effects: Some Formal Illustrations," University of Wisconsin, Institute for Research on Poverty: Discussion paper 123-72, April 1972.

Goodman, L., "Ecological Regression and the Behavior of Individuals," American Sociological Review, 1953, 18, pp. 663-664.

Goodman, L., "Some Alternatives to Ecological Correlation," American Journal of Sociology, 1959, 64, pp. 610-625.

Goodrich, R.L., "Measurement of Cognitive Change - A Dynamic Stochastic Approach," Annual Meeting of AERA, San Francisco, 1979.

Goodrich, R.L., *Stochastic Models of Human Growth*. Thesis, Division of Engineering and Applied Sciences, Harvard University, May 1978.

Goodrich, R.L., and Carnes, P.E., "Linear Stochastic System Identification from Cross-Sectional Data," *International Conference on the Analysis and Optimization of Stochastic Systems*, September 1978, Oxford, England. Proceedings to be published, Academic Press, 1979.

Goodrich, R.L., and Carnes, P.E., "Non-Stationary Linear System Identification from Cross-Sectional Data," *I.E.E.E. Trans. Automatic Control*, *AC-24*, 3, p.403, June *1979(a)*.

Goodrich, R.L., and Carnes, P.E., "On the Non-Stationary Covariance Realization Problem," *I.E.E.E. Trans. Automatic Control*, *AC-24*, 5, p.765, October *1979(b)*.

Goodrich, R.L., and Carnes, P.E., "Necessary and Sufficient Conditions for Local Identifiability," *I.E.E.E. Trans. Automatic Control*, *AC-24*, 1, p.125, February *1979(c)*.

Goodrich, R.L. and Singer, J.D., "Cognitive Change in the NDCS", Manuscript, Abt Associates Inc., Cambridge MA, 1978.

Greenhouse, S.W., and Geiser, S., "On Methods in the Analysis of Profile Data," *Psychometrika*, 1959, *24*, pp. 95-111.

Graybill, F.A., *An Introduction to Linear Statistical Models*, Vol. 1. New York: McGraw-Hill, 1961.

Gross, A.M., "Confidence Intervals for Bisquare Regression Estimates," *Journal of the American Statistical Association*, 1977, *72*, #358, pp. 341-453.

Hannan, M.T., *Aggregation and Disaggregation in Sociology*. Lexington Mass.: Lexington Books, 1971.

Hannan, M.T. and Burstein, L., "Estimation from Grouped Observations," *American Sociological Review*, 1974, *39*, pp. 374-392.

Harris, C.W. (ed.), *Problems in Measuring Change*, Madison, Wisc.: University of Wisconsin Press, 1963.

Hoaglin, D.C., and Welsch, R.E., "The Hat Matrix in Regression and ANOVA," *The American Statistician*, 1978, *21*, pp. 17-22.

Imhoff, J.P., "Computing the Distribution of Quadratic Forms in Normal Variability," *Biometricka*, 1961, *48*, No. 3-4, p. 419.

Jensen, D.R. and Solomon, H., "A Gaussian Approximation to the Distribution of a Definite Quadratic Form," J.A.S.A., 1972, 67, No. 3.

Kalman, R.E., "A New Approach to LInear Filterning and Prediction Problems," Transactions of ASME, Series D., Journal of Basic Engineering, 1960, 82, pp. 35-45.

Kalman, R.E., and Bucy, R.S., "New Results in Linear Filtering and Prediction Theory," Transactions of ASME, Series D., Journal of Basic Engineering, 1961, 83, pp. 95-108.

Kane, M.T., and Brennan, R.L., "Generalizability of Class Means," Rev. E. Research, 1977, 47, no.2, p.267.

Kendall, M., and Stuart, A., The Advanced Theory of Statistics, Vol 2, New York: Hafner Publishing Company, 1973.

Kenny, David A., "A Quasi-experimental Approach to Assessing Treatment Effects in the Nonequivalent Control Group Design," Psychological Bulletin, 1975, 82.

Knapp, T.R. "The Unit-of-Analysis Problem in Applications of Simple Correlation Analysis to Educational Research," Journal of Educational Statistics, 1977, 2, No. 3, pp. 171-186.

Lord, F.M., "Large-scale Covariance Analysis when the Control Variable is Fallible," Journal of the American Statistical Association, 1960, 55, pp. 307-321.

Lord, F.M., and Novick, M.R., Statistical Theory of Mental Test Scores. Reading, Mass.: Addison-Wesley, 1968.

Love, J.M., et al., "National Home Start Evaluation: Final Report," High Scope Educational Foundation, Ypsilanti, Michigan, 1976.

Magidson, J., "Toward a Causal Model Approach for Adjusting for Pre-existing Differences in the Non-equivalent Control Group Situation," Evaluation Quarterly, 1977, 1, 399-420.

McCall, R.B., Appelbaum, J.I., and Hogarty, P.S., Developmental Changes in Mental Performance. Monographs of the Society for Research in Child Development. University of Chicago Press, 1973, 38 (3, Serial No. 150).

Meissner, J.A., Shipman, V.C., and Gilbert, L.E., "Peabody Picture Vocabulary test," Technical Report 15, ETS-Head Start Longitudinal Study, Project Head Start, Office of Child Development, DHEW, Washington, D.C., 1972.

Mosteller, F., and Tukey, J.W., Data Analysis and Regression. Reading, Mass.: Addison-Wesley, 1977.

Porter, A.C., "The Effects of Using Fallible Variables in the Analysis of Covariance," Dissertation, University of Wisconsin, 1967. Ann Arbor, Mich.: University Microfilms (1968).

Porter, A.C., "The Effects of Using Fallible Variables in the Analysis of Covariance," Unpublished doctoral dissertation, University of Wisconsin, 1967.

Robinson, W.S., "Ecological Correlations and the Behavior of Individuals," American Sociological Review, 1950, 15, pp. 351-357.

Ruopp, R., Travers, J., Glantz, F. and Coelen, C., Children at the Center. Final Report of the National Day Care Study, Cambridge MA: Abt Associates Inc., 1979.

Shipman, U.C., McKee, J.D., and Bridgeman, B., "Disadvantaged Children and Their First School Experiences," ETS-Head Start Longitudinal Study, Princeton, September 1976.

Singer, J., "Cognitive Main Effects Analysis," National Day Care Study, 1978.

Singer, J.D., and Goodrich, R.L., "Aggregation and the Unit of Analysis in the National Day Care Study," Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 1979.

Singer, J., Affholter, D., Goodrich, R.L., "Variance Components and Reliability in the NDCS," National Day Care Study, 1978.

Strenio, J., "An Individual Growth Model Perspective for Evaluating Education Programs," Mimeograph, Cambridge, Mass.: Huron Institute, 1977.

Tukey, J.W., Exploratory Data Analysis. Reading, Mass.: Addison-Wesley, 1977.

Yule, G.U. and Kendall, M.G., An Introduction to the Theory of Statistics, Eleventh Edition. London: Griffin, 1937.

Appendix 1

APPROPRIATE WEIGHTS FOR CLASS-LEVEL REGRESSIONS

Consider the "true" linear model:

$$Y = XB + e \tag{A.1}$$

where Y is $m \times 1$, X is $m \times q$, B is $q \times 1$ and $e \sim N(0, \Sigma)$ where Σ is a diagonal $m \times m$ matrix. If we use a diagonal weight matrix W, to estimate (A.1), the weighted-class-level residual sum of squares is:

$$SS_w = (Y - X\hat{B})' \, W(Y - X\hat{B})$$

We define the weighted "hat" matrix to be:

$$H_w \overset{\Delta}{=} W^{1/2} \, X \, (X'WX)^{-1} \, X' \, W^{1/2}$$

It is then easily shown that both H_w (with rank q) and hence $(I - H_w)$ (with rank m-q) are idempotent and that $X'W^{1/2}(I - H_w) = 0$, so that

$$SS_w = (Y - XC)' W^{1/2} \, (I-H_w) \, W^{1/2} \, (Y - XC)$$

does not depend on C. Assuming the truth of (A.1), we obtain

$$SS_w = e'W^{1/2} \, (I-H_w) \, W^{1/2} \, e, \tag{A.2}$$

$$\text{or } SS_w = u' \Lambda_w u, \tag{A.3}$$

where $u \sim N(0,I)$ and

$$\Lambda_w = \Sigma^{1/2} \, W^{1/2} \, (I-H_w) \, W^{1/2} \, \Sigma^{1/2}. \tag{A.4}$$

Let λ_k, k=1,2,...m-q, be the nonvanishing eigenvalues of Λ_w. Then the form (A.3) may be rewritten:

$$SS_w \sim \sum_{k=1}^{m-q} {}_k W_k^2 \tag{A.5}$$

where W_k, $k=1,2,\ldots,m-r$, are i.i.d. $N(0,1)$. Clearly SS_w is distributed as chi-square if and only if $\lambda_k = 1$, $1<k<m-r$.

This result is used where we set $W=I$.

Now take $W = \Sigma^{-1/2}$, so that $\Lambda_w = (I-H_w)$ has $m-q$ eigenvalues equal to 0. Thus,

$$SS_w = \sum_{k=1}^{m-q} W_k^2 \sim X_{m-q}^{\;2}.$$

This result is noted where

$$\Sigma = \text{diag}\left(\frac{\sigma^2}{n_1} + \nu^2, \frac{\sigma^2}{n_2} + \nu^2, \ldots, \frac{\sigma^2}{n_m} + \nu^2\right).$$

In order to calculate W, estimates of σ^2 and ν^2 must be computed. This may be accomplished by iteratively calculating regression residuals and their variance components until convergence is obtained.

Appendix 2

BIWEIGHTED REGRESSION

Background

Ordinary least squares (OLS) regression fitting minimizes the sum of squared residuals about an estimated regression surface (hence the OLS label). For many years, OLS estimates have enjoyed widespread usage and popularity among data analysts for both the relative ease with which they can be estimated and the variety of circumstances under which they may be considered "best".

Using high-speed computers, OLS estimates can be quickly and readily obtained at a minimum of expense. Many algorithms for deriving the estimates are currently available; moreover, all sophisticated computer facilities have prepackaged programs which will produce not only the OLS estimates themselves, but a variety of additional information for evaluating the performance of the fitted surface (including estimates of standard errors, routine hypothesis tests and measures of goodness-of-fit).

One of the primary reasons that so much energy has been devoted to the derivation of techniques to produce OLS estimates to that they are known to have several desirable statistical properties. Among those of importance are:

- Unbiasedness: Under the relatively mild assumption that the error terms have an expectation of zero, OLS estimates are unbiased; that is the expected value of the OLS estimate is the population regression coefficient.
- Minimum Variance (efficiency): If one further assumes that the error terms are uncorrelated with constant variance, then OLS estimates are minimum variance among the class of linear unbiased estimators (that is the Gauss-Markov Theorem).

- Maximum Likelihood: If an even stronger assumption is made that the errors are normally distributed (which encompasses the two previous assumptions), then OLS estimators are also the maximum likelihood estimators.

If heteroscedasticity and/or dependence between observations (with known structure) are present, a "generalized" least squares solution is appropriate. This amounts ot transforming the data into a form for which the OLS solution is appropriate; then the OLS estimates on the transformed data have all the desirable properties mentioned above.

Even if heteroscedasticity is not a problem and the observations are drawn independently, however, certain conditions may exist that can prove problematic for OLS techniques. For example, a regression estimated from a small sample that is not normally distributed may give rise to estimates that are inefficient relative to nonlinear estimates. In particular, OLS estimates of location (a class of estimators including regression estimators) are susceptible to problems created by distributions that have "thicker" or "more stretched" tails than those expected under normality (Mosteller and Tukey, 1972, Ch. 10). "Contaminated" normal distributions can give rise to this problem. If a variable is distributed normally in two intermixed and otherwise indistinguishable populations, one very much smaller than the other, but also in which the variance of this variable is much larger, then the type of distribution described above will result for the population as a whole.

Under these conditions, OLS estimators are neither robust nor resistant. Resistance concerns the sensitivity, or lack thereof, of particular obtained estimates to extreme changes in the values of a small number of sampled observations. As used here, robustness refers to robustness of efficiency; this is the relative stability of the efficiency of an estimator under a variaty of assumptions concerning the distribution of the variable in the population from which the sample is drawn. Under normality, OLS estimators are efficient with respect to nonlinear estimators such as the biweight. Under the "thick-tailed" class of distributions, however, certain nonlinear regression estimators (like biweighted estimates) are more efficient than OLS estimators.

Note that there are some important problems for the estimation of a regression equation under such conditions:

- The departure from normality may not be detected, or if it is, it may be ignored. The resulting "thick-tailed" distribution is still unimodal, symmetric, and "bell-shaped."

- Usual procedures for examining residuals are unlikely to spot the problem. "Thicker" tails suggest that a higher proportion of residuals will fall into extreme ranges than would be expected under Gaussian assumptions, but these "fringe-liers" are not likely to stand out as outliers.

- Not only are OLS estimates of regression coefficients inefficient, but the usual (OLS) estimates of their sampling variances are inaccurate. Hypothesis tests based on OLS estimators, then, are suspect, and (for instance) standard errors reported in the output from standard statistical analysis computer packages (e.g., SPSS, BMD, DATATEXT) are simply wrong.

Distributions like the "contaminated" normal appear to be fairly common in social science research (see also Mosteller and Tukey, 1972, Ch. 1). In cases like the NDCS, where class- and center-level regression models may be estimated on subsamples ranging in size (roughly) from 20 to 140, the relative nonresistance and nonrobustness of OLS estimates may be problematic. Biweighting is one way to approach the problem of estimation under such conditions (Mosteller and Tukey, 1972, Ch. 10, 14-16; Andrews et al., 1972).[2]

2. Biweighted Regression

Simply stated, the biweight is a mechanism for assigning weights to observations (or cases) on the basis of the magnitudes of their deviations from a location estimate (e.g., a regression surface). Formally, the weight for observation (or case) i is:

$$w_i = \left(1 - \frac{(Y_i - y^*)^2}{(cS)^2}\right)^2 \quad \text{if } \frac{(Y_i - Y^*)^2}{(cS)^2} < 1,$$

$$w_i = 0 \text{ otherwise}$$

and yi is the observation of y for case i; y* is some estimate of location; c is an arbitrary constant (often 6 or 9); and S is an estimate of spread or scale. Typically, y* and S will be robust, resistant estimates themselves (e.g., a median or a biweighted mean; and the median absolute deviation or one-half the distance between the 25th and 75th percentiles on the sample cumulative distribution function, respectively). Clearly, the more that Y_i deviates from y*, the smaller w_i will be; at some point--determined by the values of c and S--this deviation can become so large that the weight assigned that case is zero (effectively, the observation is discarded as an outlier).

It is clear that arriving at a biweighted estimate involves iterations; the weights at one fitting of a regression, for instance, are used to establish a fitted surface (generating y*), and S (e.g., the median absolute deviation), which in turn will define new weights for the next fitting. Hopefully (and presumably) the process converges according to some preset criteria (as was the case in the NDCS); alternatively, the number of iterations to be done may be set a priori.

Classroom Process-Child Outcome Analyses

Judith D. Singer

TABLE OF CONTENTS

Page

CHAPTER ONE: OVERVIEW 324

CHAPTER TWO: DATA DESCRIPTION 327
    Adult-Focus Instrument 327
    Child-Focus Instrument 328
    Cognitive Tests 329
    Child Development Associates Checklist 330
    Policy Variables 330
    Director Qualifications 331
    Physical Environment 331
    Center Philosophy and Program Orientation 332
    Covariables 333

CHAPTER THREE: ANALYTIC ISSUES 334
    Unit of Analysis 334
        Child Level 334
        Class Level 335
        Center Level 336
    Sample Definition 336
        Adult Behavior 336
        Child Behavior 337
        Child Test Scores 338
    Approach to Analysis 338
    Construct Development 339
        Examination of the Child-Focus Instrument 339
        Examination of the Adult-Focus Instrument 342
        Summary 350

CHAPTER FOUR: CROSS-INSTRUMENT EXPLORATION: AFI-CFI LINKAGES 351
    Correlation Results 352
        Focus of Caregiver's Attention 352
        Management Behavior 355
        Social Activity 356
        Instruction and Adult Activity 356
    Additional Analyses 357

TABLE OF CONTENTS (cont.)

Page

CHAPTER FIVE: CENTER-LEVEL ANALYSIS: PROCESS-OUTCOME LINKAGE 359

- Linkage of Observations and Cognitive Test Data 360
  - Examination of Scatterplots 360
  - Correlations of Observations and Gain Scores 361
  - Development of Regression Models 367
  - Regression Results for Observational Measures 369
- Linkage of Observations and Other Center Characteristics with Cognitive Gains 373
  - Correlations of Other Center Characteristics and Cognitive Gains 374
  - Regression Results for the Other Center Characteristics 377
- Additional Validations 380

CHAPTER SIX: CONCLUSIONS 381

- Size of Effective Subgroups 381
- Focus on an Individual Child 382
- Caregiver Interactiveness 382
- Classroom Structure 383
- Child Involvement 383
- Cognitive Emphasis 383
- Individual Orientation 384
- Director Qualifications 384
- Physical Environment 385

Footnotes 386

References 387

APPENDIX A 389

CHAPTER ONE: OVERVIEW

Day care centers are complex environments in which a variety of factors may join together to influence a child's behavior and development. For example, caregivers may affect children's growth by the types and amounts of interaction they provide. The range of behavior children display may be contingent upon the kinds of activities made available to them as well as upon the size of the group during those activities. Personal qualities of the director, physical characteristics of the center and programmatic elements of the center (developmental emphasis and classroom curriculum) are clearly important to the dynamics of the day care center and thus may be associated with events a child encounters daily and with their impacts on the child.

The analysis of the relationships between observed behavior of caregivers and children (classroom process) and PSI and PPVT test scores (one important outcome for children) was undertaken with two goals in mind; first, to identify contributing factors and second, to describe the linkages between them. For convenience, this analysis is hereafter referred to as process-outcome analysis.

Each of the variables included in the process-outcome analysis has been examined separately in individual analyses with the major policy variables--staff/child ratio, group size and caregiver qualifications. The primary focus here, therefore, is not to repeat these main effects analyses, but to examine the interrelationships among these different dimensions with the intention of forming an overall picture of day care. For detailed descriptions of the separate analyses of the major components of the process-outcome analysis--the Adult-Focus Instrument (AFI), the Child-Focus Instrument (CFI), and the cognitive test scores--the reader is referred to the individual papers* in this volume or the summaries in Volume II.

The major findings of this analysis fall into two categories: those dealing with the activity of the children in care, and those dealing with the caregiver's interaction with the children. In centers with higher cognitive gains, children are more frequently involved in tasks. They move about the room with a purpose, less often wandering aimlessly. They engage in more thoughtful reflection about tasks, receive more input from adults about their activities, and cooperate more with other children. Thus, the cluster of findings on child activity shows a clear relationship between concentrated intentional child behavior and higher cognitive gains. This finding can be connected to other NDCS findings about group size in that smaller groups were associated with more of these task-oriented behaviors on the part of the children.

Higher cognitive gains were also found in centers where there was an emphasis on cognitive tasks, i.e., where the director felt that child learning was an important part of day care. Also, in centers where the director's philosophical orientation resulted in a move away from whole-group activities, and where the caregivers worked with medium- rather than large-sized groups, higher gains were found and certain caregiver behaviors predominated. In particular, caregivers in these centers interacted more with the children: They interacted socially much more than they passively observed the children; they engaged in more management of the children, giving directions and correcting mistakes more often. These findings on the interactiveness of caregivers may also be connected to the findings on group size cited in other papers. Interactiveness, too, occurs more often in centers with smaller group sizes. Thus, a clear picture emerges of centers with higher cognitive gains. They have smaller group sizes, the children are more often engaged or involved in activities, and the caregivers interact with the children more frequently in both a social and managerial role.

---

*See Goodson, 1980; Connell, 1980; Goodrich, 1980b for further information.

CHAPTER TWO: DATA DESCRIPTION

Many different types of variables were examined in the process-outcome analysis and contributed to the findings cited in the previous chapter: (a) those relating to the processes of day care--behaviors displayed by both children and caregivers; (b) those pertaining to the outcomes for children in day care; (c) those characteristics of day care centers which are presently or potentially regulatable; and finally, (d) covariables, which were to be used to set aside the effects of possibly confounding factors such as age, sex and race. A brief description of the sources of these data will be given here along with a discussion of the types of information each provided.

Adult-Focus Instrument

The Adult-Focus Instrument (AFI) includes a Physical Environment Inventory, which describes space, materials and equipment in the classroom; a Classroom Snapshot, which describes general activity patterns at a single point in time; and a Five-Minute Interaction (FMI) record, which describes the behavior of a particular caregiver in detail. The Five-Minute Interaction data are of primary interest here.

An FMI record takes the form of "sentences" identifying an actor (who), the object of the action (to whom), the nature of the action (what), and the style of the action (how). During each five-minute observation period, up to 63 sentences could be recorded, the actual number being left to the observer. For each lead teacher in the study, eighteen FMI's were typically recorded. TO WHOM codes include objects such as teacher, single child, small group (defined as 2-7 children), medium group (8-12 children) and large group (13 or more children). WHAT codes denote such behaviors as commands, instructs, praises, observes and comforts. The proportion of time each WHAT and TO WHOM code was recorded was used in analysis, rather than the raw frequency of occurrence. This was necessary because the number of sentences coded by each observer for each FMI varied. Proportions, then, reflect the proportion of total time spent in a given activity (for the WHAT codes), and the proportion of total time the caregiver's attention was focused towards a particular object (for the TO WHOM codes).

HOW codes are modifiers such as touch, non-verbal, negative, positive and happy. The HOW codes were not included in this analysis, however. This is partly because the HOW codes were only optionally noted by observers, but more importantly because the meaning of the HOW codes is likely to be highly conditional upon the WHAT and TO WHOM codes they modify. It is basically through the WHAT and TO WHOM codes, then, that information was obtained relating to caregiver behavior. A list of the AFI codes analyzed appears in Appendix A.

Child-Focus Instrument

The Child-Focus Instrument (CFI) consists of a set of codes which provide a fairly fine-grained description of child behavior in the day care setting. The CFI was administered during a 20-minute observation period divided into 100 12-second coding intervals (or 100 frames). Observers were provided with timers that clicked every 12 seconds and were instructed to record what was happening to a selected focus child at the time of each click. For each coding interval, a record was made of the specific activity in which the child was engaged and the object of the child's attention. A list of the CFI activity and attention codes used in the process-product analysis appears in Appendix A.

As with the AFI, for each 20-minute observation period, a score on each activity and object code was computed by taking the proportion of frames for which the specific activity (or object) was coded. The CFI was administered to target children in the morning hours only, during free play and teacher-directed activities. The activity codes included in the CFI can be thought of both as descriptors of classroom process and also as child behavior measures. For example, the amount of time spent in open-ended, expressive activities is indicative of classroom process, while other codes relate more directly to child behaviors (e.g., gives opinions, adds prop or idea).

Cognitive Tests

Two measures of school readiness were used in the NDCS: the Preschool Inventory (PSI) and the Peabody Picture Vocabulary Test (PPVT). The PPVT is principally a measure of receptive language functioning; the PSI is designed to measure a range of knowledge and skills including the child's knowledge of colors, shapes, sizes and spatial relationships.* Analysis has focused upon children's fall-to-spring gains on these two tests. The gain scores used were not simple differences of spring and fall scores, but were generalized gain scores, adjusted so as to avoid well-known technical problems with simple difference scores.** PSI and PPVT gain scores were used as two primary child outcome measures in the process-outcome analysis.

---

*The PSI and PPVT and their psychometric properties are described in more detail in Bache, 1980b.

**For additional information on the adjustment technique see Goodrich and Singer, 1980. These scores are essentially residual gain scores with Lord-Porter corrections for attenuation.

Child Development Associates Checklist

The Child Development Associates (CDA) Checklist is an observation instrument containing 235 items, developed by SRI International to record caregiver behavior relevant to categories of competence specified in the CDA certification system. The Checklist was filled out by observers following the 2- to 3-hour Adult-Focus Observations. An extensive analysis of the data led to identification of four CDA constructs: Child Orientation, Class Management, Resources and Environment. In addition, an overall CDA rating was assigned to each caregiver. Each of these is defined in Appendix A.

Policy Variables

Analysis in the NDCS has focused primarily upon the following regulatable characteristics.

- Number of Caregivers: the total number of caregivers assigned to each classroom.
- Group Size: the total number of children assigned to a class or to a principally responsible caregiver.
- Staff/Child Ratio: number of caregivers divided by group size.
- Caregiver Qualifications: total years of formal education, presence or absence of specialized preparation related to child care, day care experience (both prior to current job and time in current center).

Each of the above variables was examined earlier in analyses of the CFI, AFI and cognitive tests, and their results are reported elsewhere in Volume IV. The purpose of including them in the process-outcome analysis was to investigate whether or not the conclusions found in these analyses could be altered by the addition of other types of information in the model. A simple example will illustrate this idea. Analysis has shown that there is a strong association between group size and cognitive gain scores. Suppose that when information regarding caregiver behavior was included in the analysis, it was found that caregivers were more interactive in small groups and in addition, more interactive caregivers were found in classrooms where children had higher cognitive gain scores. In this situation, it would be difficult to separate out the effects of caregiver behavior and group size upon gain scores, particularly since they may all be linked in a causal chain. Examination of these variables in the context of such an inclusive analysis, therefore, allowed investigation of rival hypotheses. Appendix A contains a list of the precise variables used in analysis.

Director Qualifications

These variables pertain to the education, experience, training and degree of specialization of center directors. A list of variables examined appears in Appendix A.

Physical Environment

Information was taken from center space plans regarding the amount of space available to children. Several measures were constructed:

- Homeroom: play space designated as belonging to the target class of children.
- Class Space: center space in daily use by target class.
- Child Space: all space in center used daily by any of the target classes.

- Indoor Space: all center space dedicated to day care use.
- Outdoor Space: space designated as play area for the day care children.

Center Philosophy and Program Orientation

The philosophy of a center and its general approach to the care and education of young children can have profound effects on the dynamics of a classroom. During Phase II, data on center philosophies and curricula were gathered by means of interviews with directors, caregivers and center secretaries.* A program typology was then developed on the basis of this information. Programs were first sorted into two mutually exclusive (but not exhaustive) categories, depending on whether they emphasized the individual child or the group. Individual emphasis implied stress on independence, self-reliance, learning at the child's own pace and making the child feel important or special; group emphasis implied stress on sharing, cooperation, getting along with others and interaction between staff and children.

Beyond this broad distinction, centers were categorized according to other program emphases. The latter categories were not mutually exclusive; centers could be placed in several categories if their philosophies seemed to involve multiple emphases. These additional descriptive categories were:

- Ethnic and/or Religious Emphasis: applied to centers that stressed spiritual development or the importance of learning to appreciate one's own culture.
- Guidance Emphasis: applied to centers where staff stressed love, warmth, security, understanding or a "homey" atmosphere.

---

*Center secretaries were provided by the NDCS to each center to assist in routine data collection.

- Protective Emphasis: applied to centers that saw themselves as principally caring for children's basic needs in order to free their parents for work or other activities.
- Behavioral Supervision Emphasis: applied to centers that stressed obedience, discipline, manners and adult supervision and control.
- Cognitive Emphasis: applied to centers where staff stressed such activities as language stimulation or teaching ABC's and number concepts, and where the day care classroom was described as preparation for the grade school environment.

Covariables

These variables were to be used in regressions to control for the effects of possibly confounding factors, such as socioeconomic status and race. The intent here was not to estimate the effects of these variables, but rather to adjust for them in some manner if they were found to be associated with any of the measures. Covariables were selected for use in the process-outcome analysis on the basis of whether or not they were examined in each of the three main effects analyses. A list of the available covariables appears in Appendix A.

CHAPTER THREE: ANALYTIC ISSUES

Before delving into the actual data analysis, it is necessary to address four key concerns regarding the analysis: (1) the unit of analysis, (2) the sample definition, (3) the approach taken to analysis, and (4) construct development.

Unit of Analysis

Choice of the appropriate unit of analysis is always a key decision in any investigation involving hierarchical (nested) levels of data. Variables available for the process-outcome analysis exist at three levels of aggregation: child, class and center.* Any single time point measurement at a lower level, such as CFI or AFI data, can always be aggregated to a higher level for analysis. Generalized gain scores for the cognitive test data, however, can only be examined at the child and center levels. This is due in part to the fact that over the day care year, children within a given center frequently switched classrooms, making it difficult to create classroom level change measures without severely reducing the sample size.

Child Level

Analysis was not performed at the child level for several reasons. Among the various types of data of major interest here, only two exist at the child level: child behavior data (CFI) and cognitive test scores. Variance components analyses conducted on CFI data suggested that it is more reasonable to think of the instrument as a measure of classroom and/or center characteristics than as a measure of individual child behavior.[1] That is, the CFI should not be considered a measure of child traits, which could in turn be associated with test scores, but rather as one measure of overall classroom (or center) process. Only the cognitive test scores and the child-level covariables, therefore, are considered child-level data. However, the previously established lack of within-center/class effects for the cognitive test scores insures that aggregation bias effects are not problematic; furthermore, it suggests that the appropriate unit of analysis is the more aggregated level.* On the basis of these arguments, then, no process-outcome analyses were conducted at the child level.

---

*It is important to note that the class level and teacher level are virtually synonomous; with infrequent exceptions, the classrooms observed had only one lead teacher.

Class Level

The class level is the next level of analysis. Several main effects analyses were conducted at this level, most notably the analysis of the Child-Focus and Adult-Focus Instruments. Had the generalized cognitive change scores been available for nearly all children at the classroom level, the majority of the process-product analyses might have been conducted at this level, but classroom organizational changes were so frequent that too few children were present in the same class at both testings. Since observation data were available at the class level, however, the linkages between adult behavior and child behavior were investigated at class level. Included in this portion of the process-outcome analysis, then, are all CFI and AFI data, as well as the major policy variables and classroom-level averages of the child- and teacher-related covariables. No center-level data, including director qualifications and center philosophy, were included in this class-level analysis, however.

---

*This result and several other statistical considerations in the choice of analysis unit are discussed in more detail in Singer, Affholter, and Goodrich, 1978.

Center Level

All remaining linkages in the process-outcome analysis have been investigated at the center level. For the most part, these take the form of combined and individual associations between the PSI and PPVT generalized gain scores and adult behaviors, child behaviors, policy variables, CDA variables, director qualifications, physical environment data and center philosophy and program orientation information. In addition to the justification of analyzing gain scores at the center level based upon the relative instability of classes over the day care year, a rationale for this approach can be mounted on the basis of the requirement for unbiased hypotheses testing. Moreover, as previously cited, the absence of significant child-level covariable model (i.e., within-center or within-class effects) suggests that aggregation effects are not problematic with a center-level analysis.

Sample Definition

Once the unit of analysis issue has been treated, the next concern is to determine which data should be aggregated to the chosen levels. For example, should all caregivers observed be included in analysis--teachers and aides--or should attention be restricted only to lead teachers? The sample definition may be outlined as follows.

Adult Behavior

Data on lead teachers only, rather than on aides or on all caregivers (teachers and aides) are used in NDCS analyses of adult behavior. The aides for whom observations were made constitute neither the total population of aides in the centers under study, nor a random sample of these aides.[2] Under these circumstances, aggregating the data accross teachers and aides could lead to less meaningful variables than is desirable. Hence, only data on lead teachers is used in the process-outcome analysis.

A second concern is whether fall or spring data or both are to be analyzed. In the fall data collection, race of observer and race of caregiver were confounded to some extent; this situation was remedied by the assignment of biracial pairs of observers to teachers in the spring. In addition, previous analysis of the AFI has revealed some differences between behavior patterns in the fall and spring. The process-outcome analysis has therefore concentrated on spring data which is believed to be less influenced by observer and "start-up" effects.

Child Behavior

The Child-Focus Instrument was administered once to children during free play, once during teacher-directed activity and a third time during a randomly chosen activity type. Differences were found in CFI code frequencies across activity types, and hence care must be taken when aggregating observations over teacher-directed and free play activities. In the interest of describing child behavior in the classroom with as few variables as possible, however, the observations were combined across both activity periods for use in the process-outcome analysis. These combined variables better reflect the overall behavior patterns across all activity types during the day care day.

In addition to this analysis, a separate analysis linking CFI teacher-directed data with AFI lead teacher data was conducted. As with the AFI data, only observations made in the spring were examined. This analysis is described in Connell (1980).

Child Test Scores

Children had to have both a valid pretest and post-test for either the PSI or PPVT as well as valid CFI data to be included in analysis. It was not necessary to have valid test scores for both tests, however, and hence the sample of children for the PSI is slightly different from that for the PPVT. In addition, only children whose race was reported as white or black were examined; all children reporting race as "other" were omitted from analysis (less than 4 percent of all children).

Approach to Analysis

With the unit of analysis decisions completed and the sample definition in mind, an outline of the analytic approach may now be specified. The first stage of analysis involved within-instrument (AFI and CFI) data reduction at the class level. Due to the number of variables available from each of these sources, it was necessary to study the correlations among them to attempt to reduce the information contained in the data to as few variables as possible. A large portion of this type of examination had already been done in the main effects analysis of each instrument. Since the sample definition was slightly different in the process outcome analysis than it was in the individual main effects analysis, however, it was necessary to extend this investigation.

Once the information contained in each of the observation instruments was reduced down to a manageable number of variables, linkages between child behaviors (CFI) and adult behaviors (AFI) were examined at the class level. This is the first major aspect of the process-outcome analysis. The second portion of the analysis was the exploration of linkages among cognitive test scores, child behaviors, adult behaviors, policy variables and other regulatable center characteristics conducted at the center level. The following sections of this chapter discuss the data reduction process. The next chapter reports the linkages among instruments.

Construct Development

In the first step of the analytic process, the Child-Focus and Adult-Focus instruments were examined individually in an attempt to reduce the number of variables to a small subset of major variables. The results of this process are described below for each instrument.

Examination of the Child-Focus Instrument

The Child-Focus Instrument has been used in the NDCS to provide a fairly fine-grained description of child behavior in the day care setting. Often, however, the description it produced could be considered too detailed, in that it is very difficult to distinguish many of the fifty-four individual codes from each other. As a result, examination of the separate codes could produce a severely distorted picture of child behavior in day care. For example, the two codes CONSIDERS and ADDS PROP both reflect active, involved behavior on the part of children; however, it is very difficult for observers to differentiate between them. Classrooms that are actually characterized by high levels of this type of behavior, then, may exhibit only small to medium amounts of each individual activity. To alleviate this problem, certain codes have been combined to form macro-codes. All the codes that refer to a child receiving information, praise or other comments from adults have been grouped together under the one RECEIVES INPUT code, for example. Certain of the codes that appear under the CFI heading in Appendix A, then, are actually combinations of several individual codes (e.g., OPEN ACTIVITY, CLOSED ACTIVITY, COOPERATES). For a discussion of how these macro-codes were constructed, the reader is referred to the CFI main effects analysis (Connell, 1980).

Although this type of data reduction decreased the number of variables necessary for analysis, it did not eliminate all of the redundant information contained in the data. That is, it was possible to reduce the data set further by studying the correlations among the various CFI codes, and then grouping codes according to those relationships which made sense substantively. Table 3.1 displays the correlations among the CFI codes included in the process-outcome analysis. On the basis of the correlations contained in the top three sections of Table 3.1, three macro-codes were constructed for the CFI. Due to the relatively high correlation between CONSIDERS and ADDS PROP (.30), as well as the previously mentioned problem of differentiating between these two behaviors, the two codes were summed into the macro-code REFLECTION/INNOVATION. Once these two codes were combined, it became apparent that classrooms which were characterized by a high degree of considering, comtemplating and adding props were also characterized by much lower levels of wandering. (Note that the correlation between REFLECTION/INNOVATION and WANDERS is -.26.) To condense the information contained in these variables, another new variable was created as:

INDIFFERENCE = WANDERS - (REFLECTION/INNOVATION)

This measure can be thought of as the balance in the classroom between children wandering around as opposed to participating in thoughtful, creative, problem-solving activity.*

The third macro-code was a combination of OPEN ACTIVITY and CLOSED ACTIVITY:

CLASS STRUCTURE = OPEN ACTIVITY - CLOSED ACTIVITY.

---

*Although TASK PERSISTENCE is also highly correlated with several veriables, the actual meaning of this code is unclear, and hence it was not included in any macro-code.

Table 3.1

Child-Focus Instrument Correlations
Class Level (N=146)

	Verbal Initiative	Considers	Adds Prop	Wanders	Receives Input	Task Persistence	Non-Involvement	Moves with Purpose	Monitors Environment	Cooperates	Attention to Adult	Attention to Child	Attention to Group	Attention to Environment	Open Activity	Closed Activity	Reflection/ Innovation	Indifference	Class Structure
Verbal Initiative	1																		
Considers		1																	
Adds Prop		.30**	1																
Wanders		-.20*	-.21*	1															
Receives Input	-.18		-.10	-.17	1														
Task Persistence	.09	-.27**	-.24*	.36**	-.25**	1													
Non-Involvement	.14	.24*		-.10		-.35**	1												
Moves with Purpose	.14		.09		-.21*	.17	-.15	1											
Monitors Environment	-.12	.09	-.12		-.24*	.32**			1										
Cooperates				-.09	.22*	-.33**	.24*	-.11	-.31**	1									
Attention to Adult	-.19*		-.31**	.22*	.71**	-.19*		-.20*	-.22*	.15	1								
Attention to Child	.44**			-.16	-.25**	.21*		.10	.29**	-.22*	-.38**	1							
Attention to Group		-.32**	-.11			.12				-.12	-.19*		1						
Attention to Environment		.10	.33**	-.10	-.46**			.15			-.57**	-.29**	-.45**	1					
Open Activity		-.23*	.15	-.22*	-.46**	.27**	-.28**	.21*		-.26**	-.58**	.15		.47**	1				
Closed Activity	-.17		-.23*	-.09	.30**	-.40**	.35**	-.32**	-.36**	.43**	.35**	-.43**	.21*	-.19*	-.63**	1			
Reflection/Innovation		.78**	.83**	-.26**		-.31**	.15				-.19*		-.26**	.28**		-.16	1		
Indifference		-.55**	-.59**	.87**	-.11	.43**	-.15				.26**	-.14	.13	-.22*	-.15		-.72**	1	
Class Structure	.14	-.11	.21		-.41**	.38**	-.36**	.30**	.25**	-.39**	-.51**	.33**	-.12	.35**	.88**	-.92**			1

N.B. Only those correlations significant at the α = .15 level or better are presented.
*p<.05
**p<.01

This code describes the balance in the classroom between unstructured and structured activity. Though it is a description of child behavior, it come the closest of the CFI codes to describing the organization of the children into types of activities.

An additional note should be made here concerning the internal consistency of the CFI. Examination of Table 3.1 indicates that the information contained within the instrument can still be summarized with far fewer variables than those listed. Principal components analysis was employed to try to condense the information further; however, the inability to ascribe precise meanings to each of the components suggested that retaining the original variables would be a wise course of action. Due to the high degree of correlation among many of the variables, however, it is important to note that the use of an individual code (or macro-code) represents not only that specific behavior, but also acts as a proxy for a variety of behaviors with which it is correlated. For example, the codes RECEIVES INPUT, ATTENTION TO ADULT and CLASS STRUCTURE are all quite highly correlated; any of them can be used to characterize classes where the children receive a large amount of input from adults and hence focus their attention to adults a large proportion of the time, or to distinguish between classes in which activities tend to be closed, structured ones (such as teaching, instruction) as opposed to open-ended, free play activities. That is, the variables on the CFI tend to cluster together in several small groups which can be thought of as indicating overall patterns of child behavior.

Examination of the Adult-Focus Instrument

The Adult-Focus Instrument categorizes caregiver behavior in terms of a set of codes that are fairly self-explanatory (e.g., COMMANDS, CORRECTS, INSTRUCTS). It also records whether the caregiver's behavior was directed to other adults, a single child, a small group (defined as a group of 2 to 7 children), a medium group (8 to 12 children), or a large group (13 or more children). Although the codes contained in the AFI may reflect relatively distinct behaviors, analysis has suggested that some variables should be combined to form broader descriptions of caregiver behavior.

Table 3.2 displays the correlation matrix of the TO WHOM codes with measurements of group size, staff count and staff/child ratio from three different sources: observations, the AFI and the CFI. The TO WHOM codes reflect the proportion of time caregivers direct their attention to one of five distinct objects. Two types of correlation patterns may be found among these attention codes: those which reflect the manner in which the variables were constructed, and those which reflect actual caregiver behavior patterns. The former pattern occurs because each caregiver has only a finite amount of time which can be distributed among the five objects. Caregivers who often focus their attention to one or two of these objects can not also focus their attention towards the remaining objects, and as a consequence, many of the TO WHOM codes are negatively correlated. Whereas this type of correlation pattern is an artifact of the "limited resources" principle, the second type of correlation pattern is believed to reflect modes of caregiver behavior. For example, certain caregivers may tend to deal with children on a smaller scale and thus will spend more time with individual children and small groups and less time with medium and large groups. In this case, the caregiver chooses to spend time in this type of behavior configuration. On the basis of the correlation among these codes alone, it is difficult to distinguish between these two patterns; in conjunction with measures of classroom structure--group size, staff count and staff/child ratio--however, the picture becomes clearer.

The variable TO STAFF reflects the amount of time caregivers spend with each other as opposed to with children.

368

Table 3.2

Exploration of Dynamic Measures of Group Size
Class Level (N=118)

		TO WHOM							CHILD COUNT				STAFF COUNT				RATIO		
		To Staff	To Child	To Small Group	To Med. Group	To Large Group	Group Scale		Observed	AFI	CFI		Observed	AFI	CFI		Observed	AFI	CFI
TO WHOM	To Staff	1																	
	To Child		1																
	To Small Group		.18	1															
	To Med. Group	-.26**		-.29**	1														
	To Large Group	-.18	-.24*	-.31**	-.42**	1													
	Group Scale		-.10		-.82**	.86**	1												
CHILD COUNT	Observed			-.17	-.37**	.34**	.42**		1										
	AFI			-.37**	-.23*	.61**	.51**		.37**	1									
	CFI		-.15		-.27**	.38**	.39**		.57**	.42**	1								
STAFF COUNT	Observed			.11	-.36**	.14	.29**		.59**	.18	.26**		1						
	AFI	.43**	.		-.16					.16	.17			1					
	CFI		-.17	.15	-.27**	.18	.26**		.33**		.52**		.54**	.11	1				
RATIO	Observed			.31**		-.20*	-.10		-.37**	-.18	-.29**		.53**		.28**		1		
	AFI	.34**		.30**	-.16	-.30**	-.11		-.15	-.43**	-.23*		.21*		.10		.44**	1	
	CFI			.29**		-.17			-.18	0.31**	-.33**		.30**		.50**		.53**	.37**	1

*p<.05
**p<.01
N.B. Only those correlations significant at the α = .15 level or better are presented.

Its correlation pattern with the other TO WHOM codes is not clearly defined; although it is negatively correlated with both medium and large groups, it appears to be independent of TO CHILD and TO SMALL GROUP. Examination of the classroom structure variables (CHILD COUNT, STAFF COUNT) shows that TO STAFF appears to be a function of classroom resources; classes in which more staff are present and/or the staff/child ratio is higher, tend to have caregivers spending more time with each other.* That is, as more staff work in the classroom, more of an individual caregiver's time is spent with other staff members.

The variable TO CHILD appears to be relatively independent of all measures of classroom structure. That is, the amount of time a given caregiver spends in one-to-one interactions with individual children is fairly independent of the number of children present, the number of staff present or the resulting ratio. From the child's perspective, however, it is possible that the amount of one-to-one interaction an individual child receives is a function of these measures; more caregivers may imply more total one-to-one interaction time. Table 3.2 also indicates that caregivers who spend more time in one-to-one interaction also tend to spend more time in small groups (and less time in large groups). However, these correlations are not strong; coupled with the independence of this variable from the classroom structure measures, this suggests that it is wise to retain TO CHILD as a separate measure.

The variable TO SMALL GROUP seems to be particularly associated with classroom staff/child ratio. Moreover, it is negatively correlated with both TO MEDIUM GROUP and TO LARGE GROUP. Since it is a ratio measurement, however, and TO MEDIUM GROUP and TO LARGE GROUP appear to be more properly considered child count (or staff count) measures, it, too, was retained as a separate variable.

---

*Previous analysis has indicated that the AFI measurements of classroom structure (especially staff count) are different in kind from observed or CFI measurements. This is borne out by the low correlations between the AFI staff count and these two measures as shown in Table 3.2. For additional information see Bache, 1980a.

The data suggest that it would be wiser to combine TO MEDIUM GROUP and TO LARGE GROUP into a single variable, however. The codes are strongly negatively correlated (r = -.42), thus indicating that caregivers who spend more time in medium-sized groups tend to spend less time in large groups. While this could be an artifact of the "limited resources" principle, examination of the correlations of these two codes with the classroom structure variables suggests that this behavior may be more properly considered as due to the dynamic structure of the classroom. That is, whether a caregiver deals with a medium-sized group as opposed to a large group is affected by the child count, staff count and, to a lesser degree, the resulting staff/child ratio. As the overall classroom group size increases, caregivers tend to spend less time with medium groups and more time with large groups; indicating that the scale of the classroom (as measured by child count) determines the scale of the groups that the caregiver focuses attention upon. Subtracting these two measurements, then, creates a variable relating to the scale of the groups actually found in the classroom. Toward this end, a new variable GROUP SCALE was defined as:

GROUP SCALE = TO LARGE GROUP - TO MEDIUM GROUP.

This variable can be thought of as representing a continuum; on one end of the continuum are classes in which the caregiver spends a sizable portion of time in medium groups and little or no time in large groups, and on the opposite end are classes which are predominantly organized into large groups of children and rarely into medium-sized groups. The variable GROUP SCALE indicates where an individual class falls along this continuum. As Table 3.2 indicates, it contains most of the information reflected in the individual codes; moreover, GROUP SCALE is quite highly correlated with all measures of child count, suggesting that it is indeed reasonable to consider it a dynamic measure of group size.

Attention now shifts from the TO WHOM codes to the WHAT codes; once again, the aim is to describe caregiver behavior patterns with as few variables as necessary. Table 3.3 contains a correlation matrix of the AFI codes, grouped according to their correlation pattern. On the basis of this matrix, three AFI macro-codes were constructed--one for each of the first three blocks of variables. From the first block, as previously discussed, the variable GROUP SCALE was constructed. The next two codes that appear in the table are COMMANDS and CORRECTS; both codes reflect management behavior by caregivers. Due to the high correlation between these two codes ($r = .45$) and their similarity in content, a MANAGEMENT macro-code was defined as:

MANAGEMENT = COMMANDS + CORRECTS

Although these variables are also correlated with several of the other behavior codes, the pattern that emerged suggested that it would be wise to treat these two separately from the remaining codes; moreover, the remaining codes do not seem to be related in content to these two codes.

The next five variables displayed in Table 3.3 also form a cluster: DIRECT QUESTIONS, RESPONDS, COMFORTS, PRAISES and OBSERVES. All of these codes pertain to the social interactiveness of caregivers; the first four are aspects of an interactive caregiver, while OBSERVES relates to the noninteractive caregiver. The correlation pattern among these codes reflects the meaning one may ascribe to each of them; the first four are all positively correlated with each other (with r's ranging from .10 to .57), while OBSERVES is negatively correlated with each of them. On the basis of these correlations the macro-code SOCIAL ACTIVITY was defined as:

Table 3.3

Adult-Focus Instrument Correlations
Class Level (N=118)

	To Staff	To Child	To Small Group	To Medium Group	To Large Group	Commands	Corrects	Direct Questions	Responds	Comforts	Praises	Observes	Instructs	Adult Activity	Group Scale	Management	Social Activity
To Staff	1																
To Child		1															
To Small Group		.18	1														
To Med. Group	-.26**		-.29**	1													
To Large Group	-.18	-.24*	-.31**	-.42**	1												
Commands	-.13	.49**	-.16	.10	.11	1											
Corrects		.42**				.45**	1										
Direct Questions	-.25**	.58**		.14	-.11	.32**	.16	1									
Responds		.54**	.18	-.17	-.17		.17	.24*	1								
Comforts		.42**	.21*	-.16	-.16			.10	.28**	1							
Praises	-.17	.60**				.27**		.50**	.57**	.25**	1						
Observes	-.17	-.27**	.11		.41**	-.15	-.10	-.19*	-.34**	-.15	-.30**	1					
Instructs	-.29**			.28**				.15	.16	-.13	.26**	-.29**	1				
Adult Activity	.30**	-.39**	-.19*		-.12	-.36**	-.24*	-.31**	-.15		-.29**		-.14	1			
Group Scale		-.10		-.82**	.86**			-.15				.25**	-.11	-.11	1		
Management		.54**	-.14			.86**	.84**	.29**	.10		.18	-.15		-.35**		1	
Social Activity		.57**			-.36**	.26**	.13	.54**	.57**	.32**	.66**	-.87**	.31**	-.19*	-.23*	.23*	1

*p<.05
**p<.01
N.B. Only those correlations significant at the $\alpha = .15$ level or better are presented.

SOCIAL ACTIVITY = DIRECT QUESTIONS + RESPONDS +
COMFORTS + PRAISES - OBSERVES.

This variable may be regarded as the excess of time spent socially interacting with children as opposed to passively observing them.* The variable INSTRUCTS was excluded from this construct because it was negatively correlated with one of the component codes, COMFORTS.

The remaining AFI code, ADULT ACTIVITY, does not appear to fit into a discernable pattern, and hence will be treated separately. Although it is negatively correlated with MANAGEMENT and its component codes, the variable does not refer to caregiver behavior towards children, and hence the codes were not combined.

Just as the CFI appears to be an internally consistent measure of child behavior, the AFI also seems to provide a fairly clear and consistent profile of adult behavior. Note that in Table 3.3, the variable TO CHILD is quite strongly related to both MANAGEMENT and SOCIAL ACTIVITY; interactive caregivers tend to focus their attention to individual children more often than do noninteractive, passive caregivers. Moreover, the strength of the simple correlations among the behavior codes suggests that caregivers who are interactive display their behavior in a variety of ways which fit together into a pattern. That is, caregivers who display one of the interactive behaviors also tend not to spend a large proportion of their time passively observing children. Interactive caregivers also tend to instruct more often, although this pattern is not as clearly defined as that for the other codes. Finally, Table 3.3 shows that caregivers who spend a large proportion of their time in adult-related activity spend less of their time interacting with children (especially on an individual basis) and, as expected, more time directing their attention towards other adults.

---

*AFI main effects analyses used two alternative formulations of the social interactiveness code. For a description of the construction of these codes see Goodson, 1980. Note, however, that the correlations among these three constructs is quite high (r's from .7 to .9) and thus they may all be thought of interchangeably.

Summary

Both the Child- and Adult-Focus Instruments were analyzed individually and were found to have a great deal of internal structure. In an effort to remove redundant information from each of the instruments, several macro-codes were constructed which condensed the information contained in each of these sources.

From the CFI, it appears that classes can be classified into several distinct categories across a variety of dimensions. One dimension which classes may be distinguished by is the degree to which children are actively involved in classroom activities. Several codes describe this dimension: CONSIDERS, ADDS PROP, WANDERS and NON-INVOLVEMENT. A second dimension refers to the degree to which classes are structured by a great deal of adult involvement; codes such as RECEIVES, ATTENTION TO ADULT and ACTIVITY TYPE all reflect this dimension. The variable COOPERATES seems to represent a third distinct dimension, although it too is correlated with the degree of classroom structure.

The AFI also produced a series of dimensions by which to classify classes. The first dimension is interactiveness; this is represented by the code TO CHILD and the two macro-codes SOCIAL ACTIVITY and MANAGEMENT. The macro-codes appear to measure two relatively distinct dimensions of interactiveness; however, they are both highly correlated with the TO CHILD code. The second major dimension derived from the AFI was GROUP SCALE which reflects the size of the groups into which adults tend to structure the classroom. A third dimension is the degree to which the caregiver spends time instructing.

CHAPTER FOUR: CROSS-INSTRUMENT EXPLORATION: AFI-CFI LINKAGES

With the internal structure of each instrument more clearly defined, it is now possible to explore the relationships between adult behavior and child behavior. Some caveats are in order. First, it is important to note that although both instruments record descriptions of day care process, the profiles they provide are from two different perspectives, and derive from different days of observation. The AFI by its very nature records the overall process and structure of the day care classroom. By concentrating upon a focal caregiver, who must in turn interact with the entire class, the instrument picks up information concerning the manner in which the caregiver shapes, manipulates and structures her class. On the other hand, the CFI pays direct attention to what an individual child is doing during a twenty-minute period; even aggregated to the classroom level, the information it provides does not relate specifically to the process of day care. For the most part, it describes specific activities or behaviors, such as cooperation, giving opinions or wandering, as opposed to describing the overall tone of the class. Moreover, there is no one-to-one relationship between many of the child behavior codes and adult behavior codes; in fact, almost all the AFI WHAT codes map into the single CFI RECEIVES INPUT code. Adding to these comments the fact that observers for the AFI and CFI were present on different days implies that many relationships between AFI and CFI codes may be weak. In fact, the strongest linkages should be through the AFI TO WHOM codes and the "structural" or "process" measures from the CFI: the attention codes and the macro-code CLASS STRUCTURE. These codes describe similar classroom features across the instruments, and their relationships may be strong. Such findings would be helpful in serving to cross-validate the instruments.

Correlation Results

Table 4.1 presents the correlations between the AFI and CFI variables at the classroom level; as before, an attempt has been made to group variables according to their correlation pattern. A description of the linkages found will be presented by AFI groups, and will concentrate on those correlations significant at the .05 level or better.

Focus of Caregiver's Attention

How a caregiver distributes attention, as opposed to what she does, appears to be relatively strongly associated with child behavior. That is, how much time a caregiver spends with each of the various "objects" available, and how she structures the classroom (into small, medium or large groups) is often related to the types of behavior children display and the focus of their attention.

Classes in which caregivers spend a large proportion of their time dealing with other caregivers (frequency of TO STAFF is high) tend to be those in which children are more frequently found wandering and inattentive toward groups of people. These classes are more often organized into open-ended activities than into structured ones; moreover, the children tend not to receive input from adults as frequently. Thus, classes in which the lead teacher spends a great deal of time interacting with other adults seem to be loosely structured for the children, and many children do not appear to be involved in activities. On the other hand, TO STAFF is correlated positively with TASK PERSISTENCE implying, perhaps, that those children who are involved in a task are not interrupted, but rather can pursue tasks for longer periods of time.

One-to-one interaction with children (TO CHILD) appears to have little relationship with child behavior.

Table 4.1

AFI-CFI Cross-Instrument Correlations
Class Level (N=118)

	To Staff	To Child	To Small Group	To Medium Group	To Large Group	Group Scale	Commands	Corrects	Management	Direct Questions	Responds	Comforts	Praises	Observes	Social Activity	Instructs	Adult Activity
Monitors Environment				-.12	.11	.14				-.12							
Wanders	.21*	-.13	-.16	-.14	.28**	.25**		.15	.13	-.14		-.19*	-.13		-.12		.12
Moves with Purpose											.12		.17	-.14	.16		
Verbal Initiative		.16	.13		-.17	-.12				.16	.23*		.15		.14		-.16
Non Involvement	-.15	-.17		.11							-.13	-.25**		.17	-.16		
Receives Input	-.20*		-.18	.37**	-.14	-.29**				.16	-.12			-.10	.11	.17	
Task Persistence	.25**			-.12				.13	.11	-.12	.13			-.10			
Cooperates							.10			.20*			.15				
Attention to Adult			-.23*	.21*			.14	.14	.17	.11			.11			.18	
Attention to Child			.15		-.11		-.18	-.25**	-.25*	-.16							.11
Attn. to Environment	.17	.12	.20*	-.24*							.18					-.19*	
Attention to Group	-.26**	-.12	-.16		.18			-.11			-.18					.14	-.13
Reflection/Innovation			.12		-.22*	-.19*	-.23*	-.13	-.21*				-.11				.16
Indifference	.14		-.18	-.15	.31**	.28**	.17	.17	.20*			-.14					
Class Structure	.26**	.21*	.16	-.20*	-.18					-.10	.22*	.19*		-.11	.10	-.18	

*p<.05
**p<.01
N.B. Only those correlations significant at the ∝ = .15 level or better are presented.

The only correlation significant at the .05 level was that of TO CHILD with CLASS STRUCTURE. Adult interaction with individual children occurs more often in classes with a concentration of open-ended, expressive activities. One interpretation of this result is that one-to-one interaction occurs in all classrooms to the same degree; that is, all teachers must, at times, stop other activities and pay attention to individual children who are crying, wandering, or teasing other children. It could be that this factor is stable across classrooms of different types and varying teacher styles.

The amount of time caregivers spend focusing their attention TO SMALL GROUPS is significantly associated with two child behavior measurements. Classes organized into small groups tended to be characterized by children attending more to the environment and less to adults. This might seem anomalous since a child in a small group directed by the teacher should be attending less to the environment and more to adults. The explanation is probably that the CFI and AFI were done on different days so that there is no reason to believe that the child under observation is a member of the small group to whom the caregiver is paying attention. In fact, the class in which the child is observed may have no small groups, and even if the class is organized into small groups during the CFI observation, the child being followed may be in another group or playing alone.

Caregivers who focus their attention to medium-sized groups tend to have children who are involved in more structured types of child activities. Children in these classes pay more attention to an adult and less to the environment, and they receive more input from adults. In opposition to these findings, caregivers who focus their attention to large-sized groups tend to have more children wandering and fewer involved in thoughtful activities. The balance between these two activities, as exemplified by the code INDIFFERENCE is clearly in the direction of wandering (r = .31). Whereas children in classrooms frequently organized into medium-sized groups seemed to be attending to adult-directed tasks, children in large groups (usually the whole class) were often uninvolved in activities and wandering aimlessly about the classroom.

In sum, the AFI attention codes TO STAFF, TO MEDIUM GROUP, TO LARGE GROUP, and GROUP SCALE show clear patterns of relationships with child behaviors. Classes in which caregivers pay more attention to other staff or to large groups have more uninvolved children and the children receive less input from caregivers about their activities. Conversely, classes in which caregivers pay more attention to medium-sized groups of 8 to 12 children--effectively a subgroup of the class--have children who are involved in structured activities, receiving input from adults to whom they are paying attention.

Management Behavior

Classes in which caregivers spend more time in management behaviors (commanding and correcting) tend to be those in which the balance between wandering and active involvement of children is in favor of wandering. The children are infrequently engaged in reflective behaviors and are not paying much attention to other children. This seeming lack of child attention to tasks in highly "managed" classrooms does not mean that management behavior causes wandering or lack of thought; in fact, this is a prime example of a situation in which child behavior may influence adult behavior. That is, because the children are not involved in classroom activities, it may be necessary for caregivers to command and correct them more often to maintain order in the classroom. Alternatively, if adults spend so much time managing the class, they may not have (or make) time to encourage child involvement. Hence, it is important to realize that these correlations merely represent associations; it is not possible to infer a causal link between adult and child behavior. Note, for example, that no association was found between the degree of class structure and management behavior; commanding and correcting occur as frequently in open-ended, expressive activities as in closed, structured activities.

Social Activity

The results for the components of socially interactive behavior are not clearly defined; only a few of the component codes are related to child behavior, and no CFI codes are significantly related to the macro-code SOCIAL ACTIVITY. In general, classes in which caregivers displayed social behavior as opposed to passive observational behavior tended to be characterized by more frequent verbal initiatives on the part of children, more receipt of input from adults, less wandering, more moving with purpose, and less time spent not actively involved in a task. Moreover, these classes tended to spend more time in open, unstructured activities than in closed, structured ones. But none of these relationships are significant at the .05 level, so they must be viewed cautiously. That is, it is important not to place much faith in them as stand-alone results, but rather to view them in the context of other study results.

Instruction and Adult Activity

The correlations found between these types of adult behavior and child measures represent basic validations of each of the instruments though the size of the correlations is small. Classes in which caregivers spend more time instructing are those in which children tend to receive input from adults more often; moreover, children tend to focus their attention towards adults or groups in these classes as opposed to the environment. As expected, these classes also seem to be conducted in more structured routines than in open-ended, expressive ones. No relationships were found between the amount of instruction and child behaviors such as VERBAL INITIATIVE and REFLECTION/INNOVATION.

None of the correlations between ADULT ACTIVITY and child behaviors were statistically significant. In this case, the trends do not fit well into a composite picture of day care and, as such, must be viewed with caution.

Additional Analyses

Two supplementary sets of analyses were carried out on these data: an examination of scatterplots to search for outliers; and a variance components analysis to determine the sources of variation in the findings. To ensure that the findings were not attributable to outliers, an extensive examination of scatterplots was conducted. Classes which appeared to be atypical (either because of range or inability to fit into the overall pattern) were set aside from analysis and correlation matrices were re-estimated. The relationships (or lack thereof) found in these new correlations did not greatly differ from those in the original matrices, and hence the results described above appear to be rather stable.

Second, variance components analyses were done on the AFI and CFI variables. Results indicated that both the measures have low reliabilities. That is, the amount of systematic variation in the data attributable to various sources is small relative to the total variation. This means that there is a great deal of "noise" or random error present in these data. As the amount of noise in the data increases, the ability to detect effects decreases. Correlations attempt to link the systematic variation present in two variables; when both variables have low reliabilities (as in this situation), the ability to make these linkages is severely diminished. Thus it is not surprising that so few significant correlations were found and that those that were found were not overwhelmingly strong.

In sum, the correlation results appear to be fairly stable across classes but to contain a great deal of noise. Relationships which were statistically significant were relatively rare, but did occur in such expected areas as the AFI TO WHOM codes and the CFI attention codes. The findings thus provide a cross-validation of the two instruments but cannot give much information on more specific relationships of teacher and child behaviors.

CHAPTER FIVE: CENTER-LEVEL ANALYSIS: PROCESS-OUTCOME LINKAGE

With a clearer understanding of each of the process measures and their interrelationships in mind, attention now shifts to the linkages of adult behavior, child behavior, regulated (or regulatable) center characteristics and cognitive gain scores. As previously discussed, this analysis was conducted at the center level (see Chapter Three).

Before delving into the details of analysis, it is important to consider the meaning of these center-level variables. For measures that occur solely at the center level, such as most physical environment or center philosophy and program orientation data, the meaning of the variables is clear. Aggregated center-level cognitive gain scores also pose no interpretation problems. Indeed, if centers are weighted by the number of children included when computing a center's gain score, then the center and child analyses yield identical estimates of regression coefficients when the two models are homologous.

Process measures, however, which have been determined to be principally class-level measures, may be problematic. Variance components analyses found the CFI not to be a child-level measure of enduring child traits, but rather a measure of overall classroom process. The AFI also appears to be indicative of classroom patterns (which here are synonymous with lead teacher patterns) as opposed to center-level patterns. By aggregating across classes within a given center which may not be any more similar in kind than a random set of classes, valuable information may be lost concerning the variability of these data. Alternatively, to conduct the linkage at the class level would require the omission of test data on many children who switched classes within a center over the day care year. Although there is no way in which to preserve the class-level variability of the observation data in a center-level analysis, this loss of information seems minor compared to that encountered if two hundred children's test scores were to be omitted from analysis. Thus, the center-level aggregated CFI and AFI data may be thought of as reflecting average classroom activity patterns and styles within a given center.

The steps employed in the process-outcome linkage essentially parallel those used in the cognitive main effects analysis. Initially, two-way plots of PSI GAINS and PPVT GAINS versus each of the available regressors were examined. On the basis of these graphs, several centers were determined to be potential outliers. Weighted correlations were then computed with and without the potential outliers. Using these matrices as a guideline, regression models were constructed to predict cognitive gain scores.*

Analysis is described in two stages: first, the examination of adult behavior, child behavior, major policy variables and cognitive test scores is discussed, and then additional center characteristics (e.g., CDA ratings, director qualifications) are added to the picture.

Linkage of Observations and Cognitive Test Data

Examination of Scatterplots

The first stage in the analysis of the relationships between generalized gain scores and child and adult behavior was to examine scatterplots of PSI GAINS and PPVT GAINS versus each of the CFI and AFI variables (as obtained from the class-level analysis). These graphs suggested that CFI data bore a strong relationship to PSI gain scores, while AFI data was associated with PPVT gain scores. From the outset it is important to note that although single time point test scores on the PSI and the PPVT are highly correlated, the generalized gain scores used in all NDCS analyses are relatively independent. (At the center level, the correlation between the cognitive gain scores used in the process-outcome analysis is 0.39.) As a result, variables that are significantly correlated with one of the measures do not necessarily have to be correlated with the other measure. Although it may be desirable to find that the same variables predict both gain scores, the philosophy in this analysis has been to treat each dependent variable separately, and then to borrow strength across the tests when interpreting results.

---

*To insure that results obtained would not be affected by such variables as socioeconomic status and race, covariables were included in various stages of analysis.

In addition to suggesting that the two tests may be associated with different types of data, these graphs also showed that there were several centers which did not fit into the overall pattern for many of the dependent and independent variables. Three of these centers were the same ones set aside from the cognitive main effects analysis. For the PSI, one additional center appeared to be rather atypical; for the PPVT there were two other centers which might be considered outliers. To insure that future results would not be unduly influenced by these centers (four for the PSI and five for the PPVT) they were both included and excluded from the next stage of analysis (correlations), to determine their effect (if any) upon results.

Correlations of Observations and Gain Scores

For each generalized gain score, weighted correlation matrices were constructed both with and without these outlier centers, where the weights were proportional to the number of children whose scores were used to compute the center-level gain scores. There are two primary reasons for weighting cases by the appropriate number of children. First, since the center-level aggregations were computed with varying numbers of children, the information contained in each case had a varying level of accuracy or stability. Second, weighting alleviates problems created by undue influence of small centers with few children, whereas equal weighting tends to disproportionately "favor" those centers.

As expected, these correlation matrices indicated that the outlier centers were unduly influencing results. For example, the correlation between PSI GAINS and COOPERATES is 0.19 if all centers are included in analysis; when the four atypical centers are omitted, the correlation jumps to 0.42. These four centers fell so far away from the general pattern that they made an effect that is actually quite dramatic appear to be just barely significant. Therefore, the outlier centers were set aside from subsequent analyses, and only the results for the remaining centers will be discussed.*

Tables 5.1 and 5.2 display the weighted correlation matrices estimated for PSI GAIN and PPVT GAIN, respec= tively. Correlations are presented only for those variables which were significantly correlated with the gain scores at the .15 level or better; only those correlations significant at the .05 level or better will be discussed. In general, the correlations confirm the previously mentioned indication that the PSI is more highly associated with CFI data and the PPVT with AFI data.

Several aspects of day care classes as described by the CFI seem to be associated with higher generalized

---

*These centers were included in several biweighted analyses, and were found to receive very low weights, thus reinforcing the notion that they were distorting the overall correlational pattern.

Table 5.1

PSI GAINS WEIGHTED CORRELATIONS

CENTER LEVEL (N=53)

		PSI Gains	Group Scale	Non-Involvement	Receives Input	Task Persistence	Cooperates	Attention to Child	Reflection/ Innovation	Indifference	Class Structure	Group Size	Specialization
AFI	Group Scale	-.26	1										
CFI	Non-Involvement	.31*		1									
CFI	Receives Input	.33*	-.33*	.21	1								
CFI	Task Persistence	-.32*		-.41**	.42**	1							
CFI	Cooperates	.42**		.37**	.21	-.44**	1						
CFI	Attention to Child	-.20					-.35**	1					
CFI	Reflection/Innovation	.43**		.31*		-.45**			1				
CFI	Indifference	-.32*	.25	-.22	-.25	.51**		-.26	-.75**	1			
CFI	Class Structure	-.24		-.43**	-.45**	.41**	-.34*	.29*			1		
	Group Size	-.36**	.40**									1	
	Specialization	.25	.24	.34*	.24	-.47**	.33*				-.30*		1

*p<.05
**p<.01
N.B. Only those correlations significant at the α = .15 level or better are presented.

Table 5.2

PPVT GAINS WEIGHTED CORRELATIONS
CENTER LEVEL (N=52)

		PPVT Gains	To Child	Group Scale	Management	Social Activity	Wanders	Moves with Purpose	Indifference
AFI	To Child	.33*	1						
AFI	Group Scale	-.41**		1					
AFI	Management	.25	.56**		1				
AFI	Social Activity	.46**	.50**	-.40**		1			
CFI	Wanders	-.34*		.24	.28*		1		
CFI	Moves with Purpose	.30*				.23		1	
CFI	Indifference	-.34*		.27*	.33*		.81**		1

*p<.05
**p<.01
N.B. Only those correlations significant at the α = .15 level or better are presented.

gain scores on the PSI. Children in these centers are more frequently observed actively integrated into classroom activities as opposed to wandering aimlessly; they engage in more reflective and innovative behavior, cooperate more, and receive more input from adults.

Caregiver behavior as measured by the AFI does not appear to bear a great deal of relationship to PSI gains, as the only correlation to reach even the .15 level is that of PSI GAINS and GROUP SCALE. There is thus a suggestion that caregivers who focus their attention more often to medium-sized groups as opposed to large ones tend to be found in those centers where cognitive gains are higher. This result has greater importance when judged in conjunction with GROUP SIZE. In particular, GROUP SIZE is significantly correlated with both GROUP SCALE and PSI GAINS. The smaller the assigned group size, the higher the PSI GAINS and the more likely a caregiver is to focus her attention on medium- rather than large-sized groups. When the assigned group size is large, PSI GAINS are smaller and it seems caregivers structure the classes into larger groups (at least, they spend more time paying attention to large groups). GROUP SCALE could thus be considered an index of effective size of the groups in the classroom, and its relationship with PSI GAINS of theoretical, if not statistical, significance.

Unfortunately, there are two major anomalies in the relationship of the observational data and PSI GAINS: the strong negative correlation between TASK PERSISTENCE and PSI GAINS (r = -0.32), and the strong positive correlation between NON-INVOLVEMENT and PSI GAINS (r = 0.31). Although disturbing, these results can be explained, at least in part, by the fact that the meaning of each of these codes may be expected to vary with the activity type (free play or teacher-directed activity) in which the observation is made; when aggregated across all activities as was done here, the interpretation of the codes becomes problematic. TASK PERSISTENCE is expected to occur more in structured situations; NON-INVOLVEMENT more in free play. When these codes (but not most of the other codes) are aggregated across structured and free play situations, one must expect anomalous findings and it is important not to allow them undue significance.

While the PSI appears to be primarily associated with CFI data, the PPVT seems to be more strongly related to AFI measurements of classroom process. Caregivers who spend more of their time in medium-sized groups as opposed to large groups tend to be found in centers where children achieve higher generalized PPVT gains. As previously noted, this result refines our understanding of the meaning of the variable GROUP SIZE. In the case of the PPVT, moreover, this relationship is even more important. Recall that in the cognitive main effects analysis, the primary foundation for stating the importance of GROUP SIZE was results obtained for the PSI. Although similar patterns were found for the PPVT, the strength of the group size finding (based upon the observational measurement GROUP SIZE), was weaker. The findings for the variable GROUP SCALE, however, provide a strong basis for declaring the significance of group size for the PPVT as well.

The amount of one-to-one interaction between children and adults (TO CHILD) is also strongly related to generalized gains on the PPVT. Not only is the quantity of this individual interaction important, but also the degree of socially interactive behavior. The correlation between SOCIAL ACTIVITY and PPVT GAINS is the strongest of any correlation between an independent variable and either of the measures of cognitive gain ($r = 0.46$). Centers in which caregivers organize medium-sized groups, pay attention to individual children, and interact socially with their children are those centers with higher PPVT GAINS.

Certain child behaviors were also found to be significantly correlated with PPVT GAINS. As with the PSI, classes in which children are more actively involved in classroom activities instead of wandering around aimlessly tend to be those in which children have higher gains. In addition, children who "move with purpose" are more often found in these higher gain centers. It is interesting to note that such variables as RECEIVES INPUT, COOPERATES, and REFLECTION/INNOVATION, which were quite strongly associated with PSI GAINS, were not related to PPVT GAINS. The measures of movement about the room or involvement in tasks were associated with the more verbal measure of child gains while a more diverse set of child behaviors were related to PSI GAINS.

Development of Regression Models

The correlation patterns outlined above describe the associations between child and adult behavior and cognitive gains on an individual basis. With these relationships in mind, it is now possible to model the combined associations of these variables through multiple regression analysis.

A parsimonious approach to model construction was taken for this analysis in order to pay careful attention to problems of multicollinearity of independent variables. This approach has several advantages. Most important, when the independent variables are not truly independent (i.e., when they are multicollinear), a less parsimonious approach may yield results that are uninterpretable, whereas a more exploratory approach often leads to a series of regression equations which in combination indicate the relationships between independent and dependent variables. Second, at the center level, there are just slightly more than fifty cases; yet there are almost forty independent variables of interest.

With these dimensions, it would be next to impossible to reasonably analyze the relationships between independent and dependent variables using all independent variables simultaneously. Note, however, that these dimensions also imply that there are problems of multiplicity_in any approach to analysis. Our ability to "borrow strength" across a wide variety of dependent variables (in the different main effects analyses), helps to protect us from this threat.[3]

On the basis of correlations in Tables 5.1 and 5.2, regression models were constructed.* In the regression models, all two-way and three-way combinations of the CFI and AFI variables were tested, initially excluding those variables that, on the basis of the simple correlations were not related to gains.** Also, the major policy variables previously found to be significantly related to cognitive gains (group size, proportion of caregivers with specialization, and mean years of caregiver experience) were included. Finally, covariables were initially used to control for possible confoundings of race of children in the center and SES characterisics of the center, although as in the other cognitive analyses, they were subsequently found to be nonsignificant.

In addition to using simply weighted least squares to estimate regression surfaces, all models were also estimated using weighted-biweighted least squares. Biweighting is a robust estimation technique designed to handle data where there may be severe departures from the normality assumptions that are generally made in estimating the standard general linear model. Large changes in coefficient estimates after biweighting indicate that the results of the simple least squares regression are unstable.[4]

---

* As was done with the correlations, regressions were weighted by the appropriate number of children. In addition, centers previously determined to be outliers were not included in this analysis.

**Process variables that had nonsignificant simple correlations were subsequently entered into regression models to further investigate their behavior. Without exception, these variables remained nonsignificant.

Regression Results for Observational Measures

Several regression models were constructed for PSI GAINS, the most informative of which are presented in Table 5.3. The models reported in the table all contain at least one CFI or AFI varible that had a significant simple correlation with PSI GAINS and a significant regression coefficient whose direction of effect was identical to that of the simple correlation (or there was a good reason for the difference).

The regressions essentially confirm the correlational result: centers in which children more frequently engage in reflective behavior, cooperate with teachers and become involved in thoughtful tasks rather than wander tend to have higher gains on the PSI; in addition, children in classes that are more structured tend to have higher gains. The stability of the results for GROUP SIZE in every model indicates that the importance of group size for PSI gains is indeed independent of other measures of classroom process. The stability of the regression coefficients after biweighting further strengthens the validity of all these findings. Note, however, that this stability is due in part to the deletion of the four outlier centers. That is, in removing those centers which probably would have caused the biweighted coefficients to differ from the initial coefficients, the biweighted coefficients have become more stable.

These models were constructed with the intention of describing as tersely as possible the type of day care center which facilitates higher PSI gains. Towards this

Table 5.3

RESULTS OF WEIGHTED AND WEIGHTED-BIWEIGHTED REGRESSIONS
DEPENDENT VARIABLE: PSI GAIN SCORE*

(N=53 Centers)

Source	Independent Variables	Weighted Regression Coefficient	t	Significance of t	Biweighted Weighted Regression Coefficient	Simple Correlation	R^2
	Group Size	-0.07	-2.25	.04	-0.07	-.36	
CFI	Reflection/ Innovation	21.89	3.22	.002	22.35	.43	
CFI	Cooperates	6.58	3.11	.004	6.77	.42	.40
	Group Size	-0.09	-2.81	.008	-0.08	-.36	
CFI	Reflection/ Innovation	21.00	2.89	.007	21.41	.43	
	Specialization	.99	1.91	.07	1.09	.25	.33
	Group Size	-0.07	-2.34	.03	-0.07	-.36	
CFI	Cooperation	6.74	-3.03	.005	7.07	.42	
CFI	Indifference	-9.15	-2.19		-9.52	-.32	.33
	Group Size	-0.08	-2.78	.009	-0.08	-.36	
CFI	Reflection/ Innovation	23.58	3.36	.002	24.53	.43	
CFI	Class Structure	-3.16	-2.35	.03	-3.30	-.24	.35

*Only those AFI and CFI variables which acted as significant predictors (α = .05) appear on this table.

end, certain variables included in the models act as proxies for a whole host of variables not entered into the model but correlated with the regressors used. For example, COOPERATES not only indicates the degree to which children tend to cooperate with caregivers, but also reflects the degree to which children receive input from adults, the amount of structure in the class and also the proportion of time children spend focusing their attention towards other children. By the same principle, the variable RECEIVES INPUT, which is not included specifically as a regressor in Table 5.3, is indeed a characteristic of centers with higher PSI gains. Due to its correlation with many of the other variables, however, it was not found to be as strong a regressor as CLASS STRUCTURE or COOPERATES, for example, and as such was not explicitly entered into the regression models. Examination of the regressions in Table 5.3 in conjunction with the correlations in Table 5.1 therefore permits an overall description of center characteristics associated with higher generalized PSI gains.

The same approach was employed to construct regression models for PPVT GAINS; the results of this analysis appear in Table 5.4. As the simple correlations indicated, many aspects of caregiver behavior are associated with higher generalized gains on the PPVT, but only one CFI variable, INDIFFERENCE, is associated with gains. Centers with higher PPVT gains tend to be characterized by more one-to-one caregiver-child interaction. These caregivers spend more time in both MANAGEMENT (commanding and correcting) and SOCIAL ACTIVITY (more time interacting, and less time passively observing classroom activities). In classes with higher gains, teachers spend more time with medium-sized groups as opposed to larger ones; children tend to be more actively involved in classroom activities instead of wandering around the class.

Table 5.4

RESULTS OF WEIGHTED AND WEIGHTED-BIWEIGHTED REGRESSIONS
DEPENDENT VARIABLE: PPVT GAIN SCORES*
(N=52 Centers)

Source	Independent Variables	Weighted Regression Coefficient	t	Significance of t	Biweighted Weighted Regression Coefficient	Simple Correlation	R^2
AFI	Group Scale	-4.07	-2.02	.05	-3.88	-.41	
AFI	Social Activity	8.63	2.71	.01	8.92	.46	.33
AFI	Social Activity	10.81	3.80	.001	11.20	.46	
CFI	Indifference	-20.47	-2.70	.01	-20.44	-.34	.32
AFI	Group Scale	-4.48	-2.24	.04	-4.64	-.41	
AFI	To Child	7.11	2.39	.02	7.36	.33	
CFI	Indifference	-17.37	-2.06	.05	-17.03	-.34	.32
AFI	Group Scale	-6.02	-3.41	.002	-6.16	-.41	
AFI	Management	24.12	3.85	.001	24.49	.25	.41
AFI	Group Scale	-5.37	-2.67	.01	-5.44	-.41	
AFI	Management	14.78	2.26	.03	14.88	.25	
AFI	Social Activity	6.77	2.14	.04	6.96	.46	.35
AFI	Group Scale	-4.16	-2.22	.04	-3.98	-.41	
AFI	Management	20.49	3.30	.002	21.15	.25	
AFI	Social Activity	6.68	2.31	.03	7.25	.46	
CFI	Indifference	-24.30	-3.29	.002	-25.63	-.34	.47

*Only those AFI and CFI variables which acted as significant predictors ($\alpha = .05$) appear on this table.

The correlations in Table 5.2 suggest that it would be difficult to estimate the effects of these different AFI variables were several of them to be included in a single regression. The final model displayed in Table 5.4 shows this not to be the case, however. It seems as though the pattern of positive and negative correlations among the regressors "push and pull" against each other in such a manner that the coefficient estimates obtained in the more inclusive model are strikingly similar to those obtained in simpler models. The initial and biweighted coefficients in all models are remarkably similar, further strengthening the stability of these findings.* That is, although the predictor variables are not truly independent, it is possible to estimate their separate effects through a single model. (Note that it was not possible to include TO CHILD in this all-inclusive model because its effects and those of MANAGEMENT and SOCIAL ACTIVITY became severely attenuated. As before, however, it is important to keep in mind that even though TO CHILD is not explicitly included in most of these regression models, it _is_ included via the two AFI macro-codes with which it is correlated.)

Linkage of Observations and Other Center Characteristics with Cognitive Gains

With a clear understanding of the relationships between adult and child behavior and cognitive gains in mind, it is now possible to expand the scope of this investigation to include other center characteristics such as director qualifications, space and center philosophy. The method of investigation is essentially identical to that used for the first linkage; without restating the specific details, discussion will begin with the weighted correlations of all variables correlated with the gains at the .05 level or better.[5]

---

*As before, this stability is due in part to the deletion of the five outlier centers.

Correlations of Other Center Characteristics and Cognitive Gains

Tables 5.5 and 5.6 present the weighted correlations for these additional center characteristics and the independent variables used in earlier analyses for the PSI and PPVT, respectively. Of the many other center characteristics listed in Appendix A, only one was found to be significantly elated to PSI GAINS: COGNITIVE EMPHASIS (r = 0.38). None of the physical environment, director's qualifications or CDA variables bore a significant relationship to PSI GAINS, though two of the director's variables and one CDA measure were related at the .15 level. In fact, each of these three lesser correlations describe connections between other variables already included in the description of day care centers associated with higher PSI GAINS. For example, the director qualifications variables are both highly correlated with SPECIALIZATION (for staff in general); more qualified directors are found in centers where the staff tend to have better qualifications overall.

Many additional center characteristics were found to be quite strongly related to PPVT GAINS, however. Of the center philosophy and program orientation variables, it seems that centers with a GROUP ORIENTATION have lower gains on the average. This result is tied to the size of the groups toward which caregivers focus their attention (GROUP SCALE) and the degree of interactiveness (SOCIAL ACTIVITY). Centers which have a group emphasis tend to be organized into larger groups and caregivers tend to interact less with these groups.

Available space does not appear to be significantly related to PPVT GAINS, although there is a slight indication

Table 5.5

PSI GAINS WEIGHTED CORRELATIONS
CENTER LEVEL (N=49 TO 53)

Source[a]		PSI Gains	Group Scale	Task Persistence	Cooperates	Attention to Child	Class Structure	Specialization	Dir. Highest Degree	Dir. Specialist	Cognitive Emphasis	Child Orientation
DIR	Dir. Highest Degree	.22		-.24		-.21	-.28*	.44**	1			
DIR	Dir. Specialization	.21	.31*	-.21	.24		-.39**	.38**	.40**	1		
PHIL	Cognitive Emphasis	.38**			.26						1	
CDA	Child Orientation	.22		-.23	.32*	-.34**					.29*	1

[a] DIR = Director Qualifications, PHIL = Center Philosophy and Program Orientation, CDA = Child Development Associates Checklist.

*p<.05

**p<.01

N.B. Only those correlations significant at the $\alpha = .15$ level or better are presented.

Table 5.6

PPVT GAINS WEIGHTED CORRELATIONS
CENTER LEVEL (N=48 TO 52)

Source[a]		PPVT Gains		To Child	Group Scale	Management	Social Activity		Wanders	Indifference		Group Orientation	Individual Orientation	Cognitive Emphasis		Child Space		Resources	Class Management	Child Orientation	CDA Score		Director Specialization
PHIL	Group Orientation	-.32*			.36**	.24	-.37**		.31*			1											
PHIL	Individual Orientation	.20		.29*			.27*			-.21		-.40**	1										
PHIL	Cognitive Emphasis	.24		.29*			.27*			-.21		-.40**		1									
SPACE	Child Space	.20					.23							.28*		1							
CDA	Resources	.28*				-.30*	.28*		-.34*			-.68**	.31*	.26				1					
CDA	Class Management	.42**		.41**			.47**		-.32*	-.34*		-.44**		.33*		.30*		.54**	1				
CDA	Child Orientation	.39**		.32*			.44**		-.28*	-.20		-.57**		.33*				.52**	.72**	1			
CDA	CDA Score	.41**		.21			.45**		-.36**	-.26		-.68**		.34*		.21		.85**	.82**	.85**	1		
DIR	Director Specialization	-.23		-.36**	.29*		-.31*									-.31*			-.38**				1

[a] PHIL=Center Philosophy and Program Orientation, SPACE=Physical Environment, CDA=Child Development Associates Checklist, DIR=Director Qualifications.

*p<.05

**p<.01

N.B. Only those correlations significant at the ∝ = .15 level or better are presented.

that centers which have more space available for children tend to have higher PPVT GAINS.

Variables from the CDA Checklist appear to be strongly related to PPVT GAINS, but it is important to note that these relationships are highly correlated with the AFI variables, especially SOCIAL ACTIVITY and TO CHILD.* Caregivers who score higher on CHILD ORIENTATION, CDA SCORE and CLASSROOM MANAGEMENT are those caregivers who spend more time in social interaction and with individual children. Hence it appears that the CDA is another source of information to describe the interactiveness of caregivers. RESOURCES also appears to be associated with PPVT GAINS, implying that more toys and creative play materials in a classroom is another bit of the overall picture of quality day care.

Regression Results for Other Center Characteristics

Regression models were constructed using the additional center characteristics in conjunction with those variables previously found to be related to the gain scores. Results of these analyses appear in Table 5.7 and 5.8. Centers where directors state that there is an emphasis on cognitive development tend to have higher PSI gain scores; this result appears to be fairly independent of other behavioral and structural aspects of the center. In addition, caregivers with a stronger CHILD ORIENTATION as measured by the CDA tend to be found in centers with higher PSI gains. Although this result is not overwhelmingly strong, it does indicate that caregiver behavior influences not only PPVT gains, but PSI gains as well.

Table 5.8 further emphasizes the importance of caregiver behavior for the PPVT. Two variables from the CDA Checklist, CHILD ORIENTATION and CDA SCORE, both contribute

---

*See Goodrich, 1980b, for a more extensive discussion.

Table 5.7

RESULTS OF WEIGHTED AND WEIGHTED-BIWEIGHTED REGRESSIONS
DEPENDENT VARIABLE: PSI GAINS

Source[a]	Independent Variables	Weighted Regression Coefficient	t	p-value	Weighted-Bi-weighted Regression Coefficient	Simple Correlation	R^2
	N=49						
PHIL	Cognitive Emphasis	1.17	2.33	.03	1.36	0.38	
POL	Group Size	0.07	-2.06	.05	-0.07	0.36	
CFI	Reflection/Innovation	20.24	2.70	.01	20.39	0.43	.34
	N=49						
PHIL	Cognitive Emphasis	1.10	2.27	.03	1.19	0.38	
POL	Group Size	-0.06	-2.06	.05	-0.07	-0.36	
CFI	Reflection/Innovation	18.06	2.46	.02	19.18	0.43	
CFI	Receives Input	10.52	1.94	.06	11.96	0.33	.39
	N=49						
PHIL	Cognitive Emphasis	0.86	1.78	.09	0.98	0.38	
POL	Group Size	-0.06	-1.92	.07	-0.06	-0.36	
CFI	Reflection/Innovation	19.27	2.75	.009	18.32	0.43	
CFI	Cooperates	6.36	2.74	.009	6.62	0.42	.44
	N=53						
CDA	Child Orientation	2.08	1.71	.10	1.92	0.22	
POL	Group Size	-0.07	-2.36	.03	-0.08	-0.36	
CFI	Reflection/Innovation	24.01	3.33	.001	24.92	0.43	.32

[a] PHIL = Center Philosophy and Program Orientation, POL = Policy Variable, CFI = Child Focus Instrument, CDA = Child Development Associates Checklist.

Table 5.8

RESULTS OF WEIGHTED AND WEIGHTED-BIWEIGHTED REGRESSIONS
DEPENDENT VARIABLE: PPVT GAINS

Source[a]	Independent Variables	Weighted Regression Coefficient	t	p-value	Weighted Bi-weighted Regression Coefficient	Simple Correlation	R^2
	N=51						
CDA	CDA Score	8.38	3.51	.001	8.99	0.41	
AFI	Group Scale	-5.21	-3.30	.001	-5.26	-0.41	
AFI	Management	27.09	4.84	.001	28.08	0.25	
CFI	Indifference	-21.78	-3.10	.004	-21.43	-0.34	.55
	N=52						
CDA	Child Orientation	6.13	2.94	.006	6.66	0.39	
AFI	Group Scale	-6.60	-3.76	.001	-6.56	-0.41	
AFI	Management	17.79	2.93	.007	18.56	0.25	.39
	N=48						
PHIL	Group Orientation	-2.20	-3.18	.003	-2.23	-0.32	
AFI	Management	20.96	3.18	.003	19.92	0.25	.27

[a] CDA = Child Development Associates Checklist, AFI = Adult Focus Instrument, CFI = Child Focus Instrument, PHIL = Center Philosophy and Program Orientation.

considerable information concerning caregiver behavior to this analysis. A separate measurement of caregiver behavior confirms the earlier AFI finding that caregivers who are more interactive and tend to actively focus their attention towards children are found in centers with higher PPVT gains. Table 5.8 also shows that centers which tend to be oriented toward the group as opposed to the individual tend to have lower PPVT gains. This further refines the hypothesis that individual attention is a key aspect of centers with higher cognitive gains.

Additional Validations

The findings in this report have been subjected to extensive methodological scrutiny. Every attempt has been made to examine potential statistical pitfalls. In addition to those already described in the text of this report (such as outlier examination), several other checks were made to insure the accuracy and stability of these results.

In addition to analyzing the data from all 57 centers, the 49-center quasi-experiment and the 8-center Atlanta Public School substudy were examined separately.* The results for the 49-center experiment are essentially identical to those for the 57-center study. Coefficient estimates are remarkably stable across the two samples; estimates of standard errors are slightly smaller in the 49-center study, thus producing slightly stronger significance levels.

Most important, biweighting proved to be a very useful analytic tool for insuring that the results are not attributable to extreme, atypical cases. In conjunction with examination of scatterplots and residuals, biweighting helped to validate findings.

---

*For a discussion of the results for the APS, see Goodrich, 1980a.

CHAPTER SIX: CONCLUSIONS

The data provide clear evidence that there are discernible patterns of child and adult behavior which are associated with children's acquisition of skills and knowledge, as measured by the PSI and PPVT. There is also evidence that the center philosophy and program orientation are related to these child outcomes. However, the data do not strongly support the hypothesis that child behaviors as measured by the CFI are directly linked to caregiver behaviors.

Size of Effective Subgroups

One of the major findings of the NDCS has been that small groups are associated with better care for children. The process-outcome analysis not only supports this finding, but also provides additional refinements to our understanding of why group size is such an important dimension of quality care. The total number of children present with one or more caregivers, as measured by an actual head count, effectively determines the size of the subgroups into which caregivers organize the classroom. As the number of children assigned to a classroom increases, the size of the subgroups toward which caregivers focus their attention increases, regardless of the prevailing staff/child ratio. The size of these "effective subgroupings" is associated with a whole range of child behaviors and outcomes.

Centers in which caregivers structure the classes into medium-sized groups as opposed to large ones have higher gains on both PSI and PPVT. Moreover, there is evidence that children in these centers are more actively involved in activities, spend less time simply wandering about, are more thoughtful and reflective in their projects and receive more input from adults. When effective groupings are large, caregivers tend to stop interacting with children and begin to stand back and passively observe classroom activities. It appears, therefore, that group size does indeed influence child and adult behavior patterns, as well as cognitive gains.

Focus on an Individual Child

The structure of the day care classroom--as measured by group size, staff count and staff/child ratio--appears to have no effect upon the proportion of time caregivers spend in one-to-one interaction. Nonetheless, centers in which caregivers spend a large proportion of their time interacting with individual children tend to have higher PPVT gain scores than centers in which caregivers tend to focus their attention to groups of children. From an individual child's perspective, however, the prevailing group size may affect the amount of individualized attention received.

Caregiver Interactiveness

Interactiveness is an important dimension of quality day care. Centers in which caregivers are more interactive and orient themselves towards children tend to have higher cognitive gains, especially on the PPVT. Further, caregivers who stand back and passively observe children instead of interacting with them are found in centers with lower cognitive gains. Although the type of interaction may be either managerial (commanding and correcting) or social in nature, social interaction is the stronger predictor. In fact, the amount of social interaction bears the strongest relationship to cognitive gains of any variable examined.

Classroom Structure

The degree to which classes are organized into structured activities as opposed to open-ended, expressive ones influences a range of child and adult behaviors and outcomes. In the more structured environments children are more involved in tasks, receive more input from adults and exhibit more cooperative behavior. In the more open classes, adults respond to and comfort children more frequently, and children persist longer at tasks and pay more attention to other individual children. Both these sets of activities seem intuitively appropriate for the degree of structure in the classes and in that sense validate the observations.

Child Involvement

Children who are active and integrated into the classroom activities have higher cognitive gains on both instruments. Moreover, centers in which children spend a large proportion of their time wandering have lower gains on the average. That is, there is a distinct pattern of child behavior characterized by such behaviors as moving with purpose, contemplating and cooperating which tend to be associated with less time spent wandering as well as higher gains. Caregivers in classes where children are found wandering about tend to exhibit more management behaviors. In this situation, however, it is difficult to determine whether adult behavior influences child behavior or vice versa.

Cognitive Emphasis

Centers in which directors and other staff explicitly stated that there was an emphasis on cognitive development techniques such as language stimulation and teaching the alphabet and number concepts tended to be characterized by higher cognitive gains. This demonstrates, in effect, that centers with this emphasis have achieved their goals; if academic preparation and school readiness were program goals and if the PSI and PPVT measure these qualities, then these centers appear to have been successful.

Individual Orientation

Centers which were described as having an individual emphasis stressing the child's personal development in such areas as independence, self-reliance and self-esteem had somewhat higher PPVT gains on the average than centers which stressed children as part of the group. Moreover, caregivers in centers with an individual orientation tend to spend more time in social interaction with children whereas in centers with a group orientation, social activity between caregiver and children was low. Effective group sizes also tended to be larger in those centers which had a group emphasis. It appears, therefore, that there is indeed a center philosophy about caregiving which permeates most aspects of adult behavior. That is, when a director describes a center as having a particular philosophy or orientation, there are many recognizable facets of that orientation that emerge in both adult behavior and child outcomes.

Director Qualifications

There is little evidence to support the association of director qualifications with cognitive gains. Analysis has suggested that directors who have had courses in child development or related fields tend to be found in centers with higher PSI gains; however, the effect is reversed for the PPVT, and none of the correlations are statistically significant. As a result, the findings for this cluster of variables are essentially null. Since the director cannot be expected to serve as a composite of the caregivers observed in her center, and is limited in the insistence she can make that her caregivers and children should behave in specified way, the lack of associations in this area is not problematic.

Physical Environment

As with director qualifications, there appears to be little relationship between the amount of space available in the day care center and cognitive gains. It is not the number of square feet in a center that influences behavior, but the way in which that space is used.

FOOTNOTES

1. The results of the CFI variance components analysis appear in: J. Singer, D. Affholter and R. Goodrich, "Variance Components and the Dependability of Measures used in the National Day Care Study", Chapter 4, Abt Associates, 1978.

2. A discussion of the differences between teacher and aide behavior appears in: National Day Care Study Phase II Research Report, (Cambridge, MA: Abt Associates, 1977), Chapter 4.

3. For a more complete discussion of this approach to regression, see Goodrich and Singer, 1980.

4. Biweighting and its uses in the NDCS is discussed in more detail in Goodrich and Singer, 1980, Appendix III.

5. Once again, several outlier centers have been deleted from this analysis. In addition, note that these analyses are based upon a slightly smaller sample size due to missing data.

References

Abt Associates Inc. "National Day Care Study Phase II Research Report." Unpublished. April 1977.

Bache, W.L. "Comparing Alternative Measures of Classroom Composition." National Day Care Study Measurements and Methods. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980a.

Bache, W.L. "A Psychometric Analysis of the National Day Care Study Phase III Child Test Battery." National Day Care Study Measurements and Methods. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980b.

Connell, D.C. "The Child-Focus Observation Effects Analysis." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Goodrich, N.N. "The Atlanta Public Schools Day Care Experiment." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980a.

Goodrich, N.N. "The Effects of Day Care in Eight Atlanta Public Schools Day Care Centers." National Day Care Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980b.

Goodrich, N.N. "An Analysis of the CDA Checklist Data." National Day Care Study Measurements and Methods. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980c.

Goodrich, R.L., and Singer, J.D. "Cognitive Change in the NDCS." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Goodson, B.D. "The Adult-Focus Observation Effects Analysis." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Singer, J.D., Affholter, D.P., and Goodrich, R.L. "Variance Components and the Dependability of Measures Used in the National Day Care Study." Unpublished. 1978.

References

Abt Associates Inc. "An Overview of the Study." National Day Care Study First Annual Report 1974-1975, Volume I. Cambridge, Mass.: Abt Associates Inc., 1976.

Abt Associates Inc. "National Day Care Study Phase II Research Report." Unpublished. April 1977.

Bache, W.L. "Comparing Alternative Measures of Classroom Composition." National Day Care Study Measurements and Methods. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980a.

Goodrich, N.N. "The Effects of Day Care in Eight Atlanta Public Schools Day Care Centers." National Day Care Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980b.

Medley, D.M. and Mitzel, H.E., "Measuring Classroom Behavior by Systematic Observation." In N.L. Gage, Handbook of Research on Teaching. Chicago, Ill.: Rand McNally and Co., 1963.

Singer, J.D., Affholter, D.P., and Goodrich, R.L. "Variance Components and the Dependability of Measures Used in the National Day Care Study." Unpublished. 1978.

Stallings, J. and Broussard, D. Final Report National Day Care Cost-Effects Study: Spring 1977 (Phase III) Data Collection. Menlo Park, Calif.: SRI, 1977.

APPENDIX A

LIST OF ALL VARIABLES INCLUDED IN PROCESS-OUTCOME ANALYSIS

(1) ADULT-FOCUS INSTRUMENT

TO STAFF	Proportion of time teacher focuses attention to other staff members.
TO CHILD	Proportion of time teacher focuses attention to an individual child; i.e., one-to-one interaction.
TO SMALL GROUP	Proportion of time teacher focuses attention to a small group--defined as 2-7 children.
TO MEDIUM GROUP	Proportion of time teacher focuses attention to a medium group--defined as 8-12 children.
TO LARGE GROUP	Proportion of time teacher focuses attention to a large group--defined as 13 or more children.
COMMANDS	Self-explanatory.
CORRECTS	Self-explanatory.
DIRECT QUESTION	Proportion of time teacher poses a direct question, e.g., "What is your favorite color?"
RESPONDS	Self-explanatory.
INSTRUCTS	Self-explanatory.
ADULT ACTIVITY	Proportion of time teacher engages in self-related activity or conversation with other adults.
COMFORTS	Self-explanatory.
PRAISES	Self-explanatory.
OBSERVES	Proportion of time teacher spends listening to or observing others.
AFI CHILD COUNT	Number of children with teacher at the beginning of the FMI.
AFI STAFF COUNT	Number of staff present (including teacher) at the beginning of the FMI.

APPENDIX A (continued)

(8) CENTER PHILOSOPHY AND PROGRAM ORIENTATION (continued)

GUIDANCE EMPHASIS	As described in Chapter 2.
PROTECTIVE EMPHASIS	As described in Chapter 2.
BEHAVIORAL SUPERVISION EMPHASIS	As described in Chapter 2.
COGNITIVE EMPHASIS	As described in Chapter 2.

(9) COVARIABLES

NUMBER OF ADULTS	Average number of adults in children's homes.
NUMBER OF CHILDREN UNDER 12	Average number of children less than 12 years in children's homes.
NUMBER OF SIBLINGS	Average number of siblings.
INCOME	Average family income.
MOTHER'S EDUCATION	Mean mother's years of education.
PSI INTERVAL	Average number of days between the two tests upon which PSI GAIN is based.
PPVT INTERVAL	Average number of days between the two tests upon which PPVT GAIN is based.
FRACTION BLACK	Fraction of black children.
FRACTION MALES	Fraction of male children.
AGE	Average age of children (in months).
STAFF AGE	Average staff age.
STAFF FRACTION BLACK	Fraction of black staff.
STAFF FRACTION FEMALE	Fraction of white staff.

APPENDIX A (continued)

(6) DIRECTOR QUALIFICATIONS

DIRECTOR'S EDUCATION	Director's years of education
DIRECTOR'S HIGHEST DEGREE	Director's highest degree (ordinal scale).
DIRECTOR'S SPECIALIZATION	Indicates whether or not director had courses in subject matter related to children and child care.
DIRECTOR'S PREVIOUS DAY CARE EXPERIENCE	Director's number of years of experience in center-based day care other than present center.
DIRECTOR'S CENTER EXPERIENCE	Director's number of years in current center.

(7) PHYSICAL ENVIRONMENT

HOMEROOM	Play space designated as belonging to the target class of children (in square feet).
CLASSROOM SPACE	Center space in daily use by any of the target classes.
CHILD SPACE	All space in center used daily by any of the target classes.
INDOOR SPACE	All center space dedicated to day care use.
OUTDOOR SPACE	Space designated as play area for the day care children.

(8) CENTER PHILOSOPHY AND PROGRAM ORIENTATION

INDIVIDUAL ORIENTATION	Indicates centers oriented towards the individual.
GROUP ORIENTATION	Indicates centers oriented towards the group.
ETHNIC EMPHASIS	Indicates centers with an ethnic and/or religious emphasis.

APPENDIX A (continued)

(4) POLICY VARIABLES*

GROUP SIZE	Observed group size.
STAFF/CHILD RATIO	Ratio of observed number of staff to observed number of children.
SPECIALIZATION	Proportion of staff present who have specialized courses in subject matter related to children and child care.
PREVIOUS DAY CARE EXPERIENCE	Mean years of experience observed staff had in center-based day care other than the center in which they presently work.
CENTER EXPERIENCE	Mean number of years that staff present had worked in current center.

(5) CHILD DEVELOPMENT ASSOCIATES CHECKLIST

CDA SCORE	Overall CDA rating of caregiver competency.
CHILD ORIENTATION	The degree to which caregiver encourages cognitive and language development, a good self-concept in children, social behavior, self-help behaviors, and active play.
CLASS MANAGEMENT	The degree to which caregiver manages class activities well and emphasizes safety.
RESOURCES	The degree to which gross motor toys and creative play materials are available and arranged successfully.
ENVIRONMENT	The degree to which classroom environment is safe and sanitary.

---

*All policy variables refer to observations made of the number of staff and children present for each class during the morning hours (9:00, 10:00 and 11:00) on each of two days in each of five months in the Phase III year, and the qualifications of these observed staff.

(2) CHILD-FOCUS INSTRUMENT (continued)

TASK PERSISTENCE	Measure of task persistence--duration of longest activity during observation.
COOPERATES	Measure of the degree to which children cooperate.
ATTENTION TO ADULT	Proportion of time child focuses attention to adult.
ATTENTION TO CHILD	Proportion of time child focuses attention to another child.
ATTENTION TO ENVIRONMENT	Proportion of time child focuses attention to the environment.
ATTENTION TO GROUP	Proportion of time child focuses attention to a group.
CFI CHILD COUNT	Number of children present at the beginning of the CFI observation.
CFI STAFF COUNT	Number of staff present at the beginning of the CFI observation.
CFI RATIO	Computed as CFI STAFF COUNT divided by CFI CHILD COUNT.

(3) COGNITIVE TESTS*

PSI GAIN	Generalized gain score for the PSI.
PPVT GAIN	Generalized gain score for the PPVT (adjusted for the covariable RACE at the child level.)

---

* These gain scores were recalculated for the sample of children included in the process-outcome analysis in the same manner as that described in Goodrich and Singer, 1978.

APPENDIX A (continued)

(1) ADULT-FOCUS INSTRUMENT (continued)

AFI RATIO	Computed as AFI STAFF COUNT divided by AFI CHILD COUNT.

(2) CHILD-FOCUS INSTRUMENT

OPEN ACTIVITY	Proportion of time spent in open-ended, expressive activity.
CLOSED ACTIVITY	Proportion of time spent in closed, structured activity.
MONITORS ENVIRONMENT	Proportion of time spent monitoring environment; child's attention is obviously directed at other people or things.
CONSIDERS	Proportion of time spent considering, contemplating, tinkering; e.g., child struggles with a problem attempting to solve it.
ADDS PROP	Proportion of time child adds a different prop or new idea, i.e., variety to his/her activity.
WANDERS	Proportion of time spent wandering around center with no apparent purpose to his/her movement. Child may be sitting or standing doing nothing, looking around the area with no apparent focus.
MOVES WITH PURPOSE	Proportion of time child moves with purpose; child is going from one activity to another; evident that there is some goal to movement.
VERBAL INITIATIVE	Proportion of time child gives opinions, comments, information or states preferences.
NON-INVOLVEMENT	Proportion of time spent in no apparent task or activity.
RECEIVES INPUT	Proportion of time child receives input from adults.

The Atlanta Public Schools
Day Care Experiment

Nancy N. Goodrich

TABLE OF CONTENTS

Page

CHAPTER ONE: INTRODUCTION 399

CHAPTER TWO: EXPERIMENTAL DESIGN 402

CHAPTER THREE: IMPLEMENTATION 406

CHAPTER FOUR: ANALYSIS 409

- Twofold Nested Classification Analysis 410
- Estimation of Variance Components 412
- Conventional Analyses of Variance 418
- Regression Analyses 426

CHAPTER FIVE: SUMMARY AND CONCLUSIONS 430

References 432

CHAPTER ONE: INTRODUCTION

This report presents the design, implementation, and analysis of a randomized experiment in eight Atlanta Public School (APS) day care centers.* The purpose of the experiment was to test the hypothesis that staff/child ratio and level of staff education significantly affect caregivers' skills and behaviors, children's behaviors and children's cognitive development, as measured by the Preschool Inventory (PSI) and the Peabody Picture Vocabulary Test (PPVT) tests. The classroom-level factorial design extends over eight centers, crosses three levels of staff education with two levels of staff/child ratio and is blocked on child age (three- and four-year olds). Experimental manipulation, including random assignment of children to classes, was conducted entirely within centers.

In this paper, the results of classroom-level analyses of the effect that staff/child ratio and level of staff eduction have on children's behavior and cognitive development are summarized. The manipulations required to implement the study design also are derived and the problems encountered in implementing the study are discussed. Additional classroom-level regression analyses of the APS data that do not depend on randomized assignments and that include other policy variables (group size, number of caregivers center experience, previous day care experience) are reported in the paper "Effects of Day Care in Eight Atlanta Public Schools Day Care Centers," (N. Goodrich, 1980).

---

*We would like to thank all of the people associated with the APS day care centers who made this experiment possible, particularly Dr. Jarvis Barnes, Assistant Superintendent for Research and Evaluation, Dr. Juanita Whatley, Research Associate, Division of Research and Evaluation; Ms. Emmalean Bonds, Director of Day Care; and the center teachers and staff who so willingly made changes to implement the experimental design.

The APS centers presented a unique situation which made it possible to implement a tandomized experiment:

- The centers operated under one auspice, with similar staffing patterns, ratio requirements and funding.
- The centers included in the study served entirely black, low SES children and had only black staff in three- and four-year old classrooms.
- Most of the centers were sufficiently large to allow random assignment of children to classrooms within center.
- Levels of staff education were clearly defined because of common educational requirements corresponding to job role: in the eight study centers, directors* had a graduate education in areas pertaining to young children (MA), lead teachers had completed a two-year day care program at Atlanta Area Technical School (AAT) and aides had no day care-oriented education beyond high school (HS) except, in some cases, for a 60-hour day care training course required by the State of Georgia for caregivers who have worked in centers for three years.
- The centers operated additional day care programs for infants and older children, so that staff/child ratios could be lowered by transferring staff to these programs if necessary.
- The APS administration was willing to permit staff transfers, job redefinitions and randomized child assignments to implement a randomized factorial experimental design.

Given this homogeneous and supportive environment, it was possible to develop meaningful experimental variations in education (MA vs. AAT vs. HS) and ratio (1:5 vs. 1:10) at the classroom level, with children randomly assigned to classes within center. Since day care classrooms are usually organized into classrooms by age, the design was blocked by age.

---

*The job titles used in this report are not the same as these used by the Atlanta Public Schools. In the APS centers directors have the job title "teacher" while lead teachers are referred to as "group leaders". Aides are called aides both in the APS centers and in this report.

In the remainder of this report, the manipulations required to effect the design are described, problems with design implementation are discussed, and results from the data analysis are presented. It is shown that although there were significant classroom effects for PSI (but not PPVT) generalized gain scores, little of this variance could be attributed to either staff/child ratio or level of staff education.

CHAPTER TWO: EXPERIMENTAL DESIGN

The APS study is based on a two-way factorial design that crosses three levels of staff education and two levels of staff child ratio. Treatments were defined for 29* classrooms distributed over eight relatively homogeous centers. Children were randomly assigned to classes within center according to age (three- and four-year olds).

Although the APS centers are quite similar in many respects, it is possible that center differences contributed to differences in children's cognitive development. Such an effect would preclude conventional two-way analyses of variance. (Adjustments for center-level effects would have to be made in the analyses.) The existence of a center effect is testable through a variance component analysis that compares within-center to across-center variations. Although the statistical power of these analyses is not great, it is sufficient to detect medium to large differences (1-1/2 to 2 points) in children's PSI and PPVT gain scores.

The basic APS design is depicted in Figure 2.1. Since the experimental ratio differences were obtained primarily by varying the number of caregivers in a classroom, staff/child ratio and number of caregivers are confounded in the design. To the extent possible, two other major policy variables--group size and years of experience--are balanced across cells. In addition, no two cells contain the same combination of centers. A detailed representation of the design showing classes, numbers of staff and children and average years of experience is presented in Figure 2.2.

---

*An additional classroom was added in one center after randomization was effected. This classroom was included in analyses that were not based in the randomized portion of the design.

Figure 2.1

APS DESIGN

	Educational Level			
Staff-Child Ratio	H	M	L	
1:5	2 classes	2 classes	2 classes	Three-Year-Olds
1:10	2 classes	2 classes	2 classes	
1:5	2 classes	5 classes	4 classes	Four-Year-Olds
1:10	2 classes	2 classes	2 classes	

Although most children spent between seven and ten hours a day at the APS centers, the study focused only on the core morning hours (9:00 a.m. to 12:00 p.m.). It is during this time that the children actively participate in a day care program. Outside these hours children are primarily involved in arrival, departure, mealtime or naptime--situations in which it is difficult, if not impossible, to maintain classroom distinctions and treatment contrasts.

The three education levels (MA, AAT, HS) included in the design were obtained by temporarily redefining job roles for some staff members. Directors, who usually assisted in all classes as needed, were responsible for one (high education) classroom. Ten aides were promoted to

Figure 2.2

DETAILED APS DESIGN

Age	Ratio	Education High (MA)	# Staff	Avg. Yrs. Exp.	Grp. Size	# Tgt. Kids	Education Medium (AAT)	# Staff	Avg. Yrs. Exp.	Grp. Size	# Tgt. Kids	Education Low (HS)	# Staff	Avg. Yrs. Exp.	Grp. Size	# Tgt. Kids	Total Tgt. Kids
3's	High	Center A, class 1	2	7*	12	5	Center F, class 2	2	2.5	11	9	Center D, class 3	2	7*	11	6	
		Center G, class 1	2	4.5	12	12	Center B, class 2	3	9.3	15	9	Center G, class 1	4	4*	21	16	57
3's	Low	Center C, class 1	1	5	10	9	Center E, class 2	2	9	20	18	Center C, class 3	2	5	20	18	
		Center E, class 1	2	7.5*	20	17	Center A, class 2	1	7	11	4	Center H, class 3	1	3	10	7	73
Total Target 3's						43					40					47	130
4's	High	Center H, class 1	2	5.5	9	6	Center A, class 3	4	4*	22	9	Center F, class 4	4	2.5*	21	20	
		Center B, class 1	4	6.5	19	11	Center C, class 2	2	5*	11	11	Center C, class 4	2	5	11	11	
							Center G, class 2	3	5.7	16	16	Center E, class 5	4	5.5*	20	16	
							Center D, class 2	2	5	11	6	Center A, class 4	2	-*	11	6	
							Center E, class 3	4	5.5*	19	14						126
4's	Low	Center F, class 1	2	5	21	13	Center E, class 4	2	6	20	13	Center D, class 4	1	7	11	7	
		Center D, class 1	1	7	11	6	Center H, class 2	1	-*	9	5	Center G, class 3	1	7	9	9	53
Total Target 4's						36					74					69	179
Total Target Kids						79					114					116	309

*Data missing for some staff

acting lead teacher (low education) positions for the duration of the study year (October 1976 - June 1977), with appropriate salary increases paid by the study. In the remaining (medium education) classrooms, lead teachers continued in their previous job roles. Since 18 new lead teacher positions were created for the study, a corresponding number of former lead teachers were asked to act as aides in either high or medium education classrooms. These staff continued to receive lead teacher salaries.

Two staff child ratio contrasts--1:5 (high) and 1:10 (low)--also were created for the study. These were obtained by assigning approximately ten, fifteen or twenty children and one to four caregivers to each class. Since the APS centers operated at an average ratio of 1:7, few staff reassignments to other programs were necessary.

To obtain appropriate ratio contrasts by age, the birth date for determining three- and four-year-old classes varied by center. Children born between January 1, 1972 and October 1, 1972 were assigned to four-year-old classrooms, children born between January 1, 1973 and December 31, 1973 were assigned to three-year-old classrooms, and children born between October 1, 1972 and December 31, 1972 were assigned to either three- or four-year-old classrooms as needed to effect the ratio contrasts. Although this resulted in uneven age distinctions across centers, there was no systematic age bias.

CHAPTER THREE: IMPLEMENTATION

Throughout the study year, classrooms were periodically observed and rostered to provide information about the on-going day care process. This data base also provided information about the extent to which the APS design was implemented during the year. If the data indicated that the specified treatment was not delivered to one or more classes, it was important to revise the design to reflect the true situation before the effects of staff/child ratio and staff education were tested. Based on these data, treatment designations were changed for several classes before analyses were undertaken. The revised design is shown in Figure 3.1.

Level of education was changed from "medium" to "low" for one classroom because initial information about the lead teacher's education was incorrect. Since level of education is defined by the type of education obtained by the lead teacher, assignments to education levels were based on independent knowledge about the type of schooling obtained (MA, AAT, HS) rather than on observed years of education or highest degree. Nonetheless, there are differences in the averages of the two observed education variables across levels which tend to verify that education contrasts were implemented.

Information obtained informally from center staff indicated that the lead teachers designated by the study were in charge of their assigned classrooms during the year. However, in two "high" level of education classes (centers F and H), the directors were reported to have spent a minimal amount of time actually working with the children in their assigned classes. This fact is reflected in the low observed education figures for these classes. Thus, there is evidence that the classes could have been reassigned to a "medium" level of education. They were retained as "high" primarily

Figure 3.1

DETAILED APS DESIGN (REVISED)

Age	Ratio	Education High (MA)	Obs. Ratio	Obs. Educ	Obs. High Degree	Education Medium (AAT)	Obs. Ratio	Obs. Educ	Obs. High Degree	Education Low (HS)	Obs. Ratio	Obs. Educ.	Obs. High Degree	Grp. Avg. Ratio
3's	High	Center G, class 1	.200	15.22	2.26	Center F, class 2	.179	13.00	1.48	Center D, class 3	.179	13.61	1.14	.180 (1:5.56)
						Center B, class 2	.161	13.46	1.46	Center H, class 1	.181	12.99	1.31	
3's	Low	Center C, class 1	.113	14.46	2.02	Center E, class 2	.145	12.52	1.56	Center C, class 3	.149	13.96	1.62	.136 (1:7.35)
		Center E, class 1	.118	14.20	2.15	Center A, class 2	.131	14.73	1.80	Center F, class 3	.142	12.38	.80	
		Center A, class 1	.157	14.74	1.86									
Avg. Educ. 3's				14.66	2.07			13.43	1.58			13.24	1.22	130
4's	High	Center B, class 1	.190	15.72	1.86	Center A, class 3	.207	14.30	1.54	Center F, class 4	.162	13.67	1.25	.190 (1:5.27)
		Center F, class 1	.163	12.61	1.37	Center G, class 2	.169	14.81	1.96	Center E, class 5	.193	12.36	1.15	
						Center E, class 3	.189	13.50	1.72	Center A, class 4	.199	13.69	1.16	
										Center H, class 2	.236	13.22	1.50	
4's	Low	Center D, class 1	.111	14.77	1.95	Center E, class 4	.132	13.77	1.90	Center D, class 4	.144	13.35	1.22	.134 (1:7.4)
		Center H, class 1	.131	12.66	1.85	Center C, class 2	.124	14.20	1.90	Center G, class 3	.128	13.17	1.11	
						Center D, class 2	.151	13.94	1.80	Center C, class 4	.155	12.99	1.45	
Avg. Educ. 4's				13.86	1.76			14.09	1.82			13.18	1.26	
Avg. Educ. Across Age Gp				14.26	1.92			13.83	1.72			13.20	1.25	

*Data missing for some staff

because the directors continued to plan the classes, even though they did not participate to the desired degree.

Nine classes were assigned to new ratio levels based on the observed ratio data. Consequently, two pairs of treatment cells include the same combination of centers. While this situation results in an undesirable confounding of center and treatment for the four affected cells, analyses based on the implemented design will not be misleading if a systematic center effect does not exist.

The ratio contrasts actually implemented in the APS centers are not as great as specified in the design. The average "high" ratio was 1:5.38 and varied from 1:4.24 to 1:6.21, while the average "low" ratio was 1:7.41, ranging from 1:6.37 to 1:9.01. While the ratio level differences are extremely significant ($p<.001$), the contrasts must be interpreted as being 1:5 vs. 1:7 rather than 1:5 vs. 1:10.

There are two probable reasons for the less distinct ratio contrasts. First, the ratio observations often included a free-play situation in which children from several classes were merged on a playground or in one large area. Ratios for all classes would be approximately the same in this situation, tending to regress toward the mean for the center. Also, several centers were assigned extra student interns or other staff after the APS study began. These staff generally were used in support activities not directly involved with teaching (toileting children, setting up activities or lunch, etc.). It is likely that the extra staff assisted more often in low ratio classes, effectively increasing both observed and actual ratios.

CHAPTER FOUR: ANALYSIS

Three types of analyses were undertaken based on the APS study design:

1. Analyses that examined a subset of randomized classes within center to determine
   (a) the existence of a center effect,
   (b) the existence of a classroom effect, and
   (c) the variance components associated with each.
2. Analyses of variance based on the revised design to determine the effects of level of staff/child ratio, level of staff education and the interaction between the two independent variables.
3. Regression analyses which used a continuous observed staff/child ratio variable (see Bache, 1980) and the three-level staff education design variable.

Each of these analyses represents a trade-off between statistical power and design validity. For this reason, multiple analysis techniques were used to examine thoroughly the existence of a treatment effect.

Dependent variables for the analyses were children's PSI and PPVT generalized gain scores (Goodrich and Singer, 1980). The analyses of variance also examined caregivers' and children's classroom behavior as observed during the spring of 1977. Caregiver behaviors were recorded using the Adult Focus Observation Instrument (Goodson, 1980); caregiver skills were measured through a checklist based on the Child Development Associates (CDA) credentialing system (N. Goodrich, 1980); and children's behaviors were indicated on the Child Focus Observation Instrument (Connell, 1980).

Two-fold Nested Classification Analysis

Although the APS experimental design is basically a 2 x 3 factorial, the design extends over several centers, and experimental manipulation is entirely within center. The most statistically powerful analysis for treatment effects is the conventional two-way ANOVA, but if there are center effects on cognitive gains (except through classroom level factors), the design is not appropriate. Center effects may arise because of differences attributable to the center itself. In addition, since children were randomly assigned to classes only within center, there may remain systematic differences in the children served by centers. A two-fold nested classification analysis (Graybill, 1961) avoids these technical problems but does so at the expense of reduced statistical power.

The statistical model is the random effects model

$$Y_{ijk} = \mu + \alpha_i + \beta_{ij} + e_{ijk},$$

where effects in the model are given by

μ = grand mean
α_i = center effect
β_{ij} = classroom effect
e_{ijr} = child/error effect

while the index structure and experimental design are specified by

i = center index

j = classroom-within-center index

k = child/error-within-classroom-within-center-index

K_{ij} = number of children in classroom j (within center i)

$K_i = \sum_j K_{ij}$ = number of children in center i

$K = \sum_i K_i$ = number of children (total)

J_i = number of classes in center i

$J = \sum_i J_i$ = number of classes (total)

I = number of centers

The object in analyzing the model is not to estimate fixed classroom or center effects but rather to estimate the components of variance associated with these sources. Results are used to inform subsequent fixed effect analyses.

The two-fold nested design includes only those 24 classrooms that were actually involved in randomized assignments of children (see Table 4.1). Differences within center in either three-year old or four-year old designs are attributable entirely to classroom-within-center effects among classrooms in the randomized portion of the design. Sources of these effects include:

1. experimental manipulation of design factors;
2. other (incompletely balanced) classroom composition factors;
3. caregiver effects unrelated to level of education; and
4. experimental disruption.

Table 4.1

TWO-FOLD NESTED CLASSIFICATION
EXPERIMENTAL DESIGN

Center	Three-Year Olds	Four-Year Olds
A	HL vs. ML	MH vs. LH
C	HL vs. LL	ML vs. LL
D		HL vs. ML vs. LL
E	HL vs. ML	MH vs. ML vs. LH
F	MH vs. LL	HH vs. LH
G		MH vs. LL
H		HL vs. LH

This design indicates that while level of education differences were implemented for every center within age group, staff/child ratio differences existed in only four of the seven centers (E, F, G, H) the within age group. Thus statistical power to test for within-center staff-child ratio effects is not high.

Estimation of Variance Components

The ANOVA displayed in Table 4.2, was used to estimate variance components for the APS design. Results from these analyses are shown in Tables 4.3 and 4.4. F-tests and significance levels in Tables 4.3 and 4.4 were computed as follows.

The test for the hypothesis that a center effect exists ($\sigma^2>0$) is constructed in terms of the statistic I_{ms}/J_{ms} (see Table 4.2), assumed to be distributed as $F_{(I-1, J-I)}$ under the null hypothesis $\sigma^2 = 0$. In the

Table 4.2

ANALYSIS OF VARIANCE FOR TWO-FOLD NESTED MODEL

(Symbols are defined as in the random effects model)

	DF	SS	MS	EMS
Centers	I-1	$\sum_{ijk} (Y_{i..}-Y_{...})^2$	I_{ms}	$\sigma_e^{\ 2} + g_1\sigma_\beta^2 + g_2\sigma_\alpha^2$
Classrooms	J-I	$\sum_{ijk} (Y_{ij.}-Y_{i..})^2$	J_{ms}	$\sigma_e^{\ 2} + g_o\sigma_\beta^2$
Children	K-J	$\sum_{ijk} (Y_{ijk}-Y_{ij.})^2$	K_{ms}	$\sigma_e^{\ 2}$

$$g_o = (K - \sum_i \frac{\sum_j K_{ij}^2}{K_i})/J-I)$$

$$g_1 = {}_{ij} (\frac{K_{ij}^2}{K_i} - \frac{K_{ij}^2}{K})/(I-1)$$

$$g_2 = (K - \frac{\sum_i K_i^2}{K})/(I-1)$$

Table 4.3

VARIANCE COMPONENTS (VC) ANALYSIS OF PSI GENERALIZED GAIL SCORES

Three-Year Olds

	DF	SS	MS	VC	F	P
Centers	3	292.70	97.57	3.78	3.10	.15
Classrooms	4	126.04	31.51	2.67	2.74	.04
Children	59	678.42	11.50	11.50	--	--

$g_o = 7.49$ $g_1 = 8.85$ $g_2 = 15.91$

Four-Year Olds

	DF	SS	MS	VC	F	P
Centers	6	244.47	40.74	.31	1.23	.37
Classrooms	9	299.00	33.22	2.14	2.50	.01
Children	137	1820.54	13.29	13.29	--	--

$g_o = 9.33$ $g_1 = 9.70$ $g_2 = 21.26$

Table 4.4

VARIANCE COMPONENTS (VC) ANALYSIS OF PPVT GENERALIZED GAIN SCORES

Three-Year Olds

	DF	SS	MS	VC	F	P
Centers	3	945.40	315.13	17.59	6.07	.06
Classrooms	4	207.63	51.91	.32	1.05	.39
Children	55	2730.26	49.64	49.64	--	--

$g_o = 7.13$ $g_1 = 8.20$ $g_2 = 14.94$

Four-Year Olds

	DF	SS	MS	VC	F	P
Centers	6	517.89	86.32	(-)	.87	.55
Classrooms	9	891.07	99.01	2.33	1.27	.26
Children	132	10269.17	77.80	77.80	--	--

$g_o = 9.11$ $g_1 = 9.22$ $g_2 = 20.51$

(-) indicates negative estimate for VC.

unbalanced case considered, this test is approximate, but may be appropriate when g_o and g_1 (the parametrs modifying the variance components in Table 4.2) are nearly equal. The test is identical to the fixed effect ANOVA test for a center effect, using the class as unit of analysis and weighting by children per class.

The test for the hypothesis that a classroom effect exists ($\sigma^2>0$) is constructed from the statistic J_{ms}/K_{ms}, assumed to be distributed as F(J-I, K-J) under the null hypothesis $\sigma^2=0$. It is equivalent to the fixed effect ANOVA test for a classroom effect using within-center deviations from the mean.

Results for the PSI (Table 4.3) indicate that for both three- and four-year olds the hypothesis of no classroom effect is rejected at .05 level, while the hypothesis of no center effect is accepted. However, the center effect for three-year olds is nearly significant and is potentially problematic for later analyses. It was determined to be caused primarily by one outlier center. When this center is dropped (Table 4.5), the classroom effect is relatively unchanged while the EMS estimate of the center component of variance becomes negative. This indicates that if the outlier center is dropped, conventional ANOVA's can be performed with little or no danger of disruptive center effects.

Table 4.5

VARIANCE COMPONENTS (VC) ANALYSIS OF PSI GENERALIZED GAIN SCORES

(Outlier Center Excluded)

Three-Year Olds

	DF	SS	MS	VC	F	P
Centers	2	27.31	13.65	(-)	.36	.72
Classrooms	3	112.86	37.62	3.28	3.02	.04
Children	45	560.93	12.46	12.46	--	--

$g_o = 7.68$ $g_1 = 9.79$ $g_2 = 17.78$

Four-Year Olds

	DF	SS	MS	VC	F	P
Centers	5	70.41	14.08	(-)	.38	.35
Classrooms	8	298.93	37.37	2.58	2.77	.01
Children	119	1608.14	13.51	13.51	--	--

$g_o = 9.25$ $g_1 = 9.88$ $g_2 = 22.03$

(-) indicates negative estimate for VC.

For the PPVT, the classroom level effect is nonsignificant, but there is a center effect for three-year olds, significant at the level p=.06 (see Table 4.6). This effect is not removed by dropping the outlier center. However, since it is evident that the experimental manipulation did not create significant classroom level variance for the PPVT in either age group, it is unlikely that the small center effect for three-year olds will invalidate results from pooled analyses. For the sake of completeness, ANOVA and regression results are presented for the PPVT, but they only serve to corroborate the lack of findings.

Conventional Analyses of Variance

Analyses of variance (unbalanced) were used to examine the effects of the design variables on children's cognitive development (PSI, PPVT), caregiver behaviors (AFI), caregiver skills (CDA) and children's behaviors (CFI). The analyses were performed both with and without the outlier center. Results for PSI and PPVT gain scores are shown in Tables 4.7 to 4.11; results for CFI, AFI, and CDA variables are summarized in Table 4.12.

For several reasons, these ANOVAs are of potentially greater statistical power than the nested design considered earlier. First, they include all 30 classes rather than the subset of 24 randomized classes. Second, dependent measures are not corrected for center effects, so that the variance of independent measures includes existing (non-experimental) differences among centers. Finally, tests for specific main effects (e.g., staff-child ratio) use only one degree of freedom for the hypothesis. The ANOVAs presented were performed at the classroom level, weighted by number of tested children; unweighted classroom-level and child-level ANOVAs replicated the results and are not presented here.

Table 4.6

VARIANCE COMPONENTS (VC) ANALYSIS OF PPVT GENERALIZED GAIN SCORES

(Outlier Center Excluded)

Three-Year Olds

	DF	SS	MS	VC	F	P
Centers	2	761.37	380.68	29.20	8.80	.06
Classrooms	3	129.74	43.25	(-)	.85	.47
Children	43	2188.61	50.90	50.90	--	--

$g_o = 7.37$ $g_1 = 9.36$ $g_2 = 11.31$

Four-Year Olds

	DF	SS	MS	VC	F	P
Centers	5	342.15	68.43	(-)	.69	.64
Classrooms	8	793.08	99.14	2.19	1.25	.28
Children	116	9182.63	79.16	79.16	--	--

$g_o = 9.13$ $g_1 = 9.48$ $g_2 = 22.81$

(-) indicates negative estimate for VC.

In Table 4.7, where all centers are included, there is a significant (P=.03) effect for staff-child ratio for three-year-olds but no other significant effect. Since previous findings indicated that a small center effect was introduced by an outlier center, the ANOVA was repeated omitting that center (Table 4.8). The results show that without center C, the three-year-old staff-child ratio effect remains nearly unchanged, while the nonsignificant ratio effect for four-year-olds become even less significant.

Table 4.9 and 4.10 display results for the PPVT. There are no significant effects when all centers are included. When the outlier center is omitted, however, a significant (p=.02) staff/child ratio effect emerges for four-year olds. Slightly more statistically powerful analysis across age groups that exclude the outlier center (Table 4.11) show no significant effects for the design variables on either PSI or PPVT gain scores. Thus there are hints but no strong evidence that staff/child ratio affects cognitive gain scores.

The fact that the variance component analysis showed a strong classroom effect for both age groups on the PSI while the ANOVA's that considered the two design factors showed only a weak staff-child ratio effect for three-year-olds suggests that the within-center variance is attributable to one or more other factors. Possible alternative factors include group size, teacher variables such as previous experience, competence or general personality, and the disruptive effects of the experiment. Given the results of subsequent analyses that replicated the 57-center study, (N. Goodrich, 1980b), it is probable that a major source of classroom level variance for the PSI is group size.

Table 4.7

ANOVA RESULTS FOR PSI (WEIGHTED)

Three-Year Olds

	SS	DF	MS	F	Sig.
Main Effects	45.92	3	15.31	3.90	.06
Staff-Child Ratio	27.46	1	27.46	7.00	.03
Level of Education	14.74	2	7.37	1.88	.22
2-Way Int.	13.44	2	6.72	1.71	.25
Residual	27.42	7	3.92	--	--

Four-Year Olds

	SS	DF	MS	F	Sig.
Main Effects	24.12	3	8.04	2.22	.14
Staff-Child Ratio	5.33	1	5.33	1.48	.25
Level of Education	16.89	2	8.44	2.33	.14
2-Way/Int.	6.70	2	3.35	.93	.42
Residual	39.72	11	3.61	--	--

Table 4.8

ANOVA RESULTS FOR PSI (WEIGHTED)

(Outlier Center Excluded)

Three-Year Olds

	SS	DF	MS	F	Sig.
Main Effects	20.42	3	6.81	3.94	.09
Staff-Child Ratio	15.28	1	15.28	8.85	.03
Level of Education	6.62	2	3.31	1.92	.24
2-Way Int.	13.50	2	6.75	3.91	.09
Residual	8.63	5	1.73	--	--

Four-Year Olds

	SS	DF	MS	F	Sig.
Main Effects	15.42	3	5.14	1.69	.24
Staff-Child Ratio	.32	1	.32	.10	.76
Level of Education	14.96	2	7.48	2.46	.14
2-Way/Int.	6.30	2	3.15	1.03	.40
Residual	27.41	9	3.04	--	--

Table 4.9

ANOVA RESULTS FOR PPVT (WEIGHTED)

Three-Year Olds

	SS	DF	MS	F	Sig.
Main Effects	31.18	3	10.39	.51	.69
Staff-Child Ratio	1.16	1	1.16	.06	.81
Level of Education	30.66	2	15.33	.75	.51
2-Way Int.	20.54	2	10.27	.50	.63
Residual	143.20	7	20.46	--	--

Four-Year Olds

	SS	DF	MS	F	Sig.
Main Effects	44.15	3	14.72	1.50	.27
Staff-Child Ratio	17.57	1	17.57	1.79	.21
Level of Education	25.58	2	12.79	1.31	.31
2-Way Int.	14.12	2	7.06	.72	.51
Residual	107.64	11	9.79	--	--

Table 4.10

ANOVA RESULTS FOR PSI ACROSS AGE GROUPS

(Outlier Center Excluded)

Three-Year Olds

	SS	DF	MS	F	Sig.
Main Effects	60.16	3	20.05	1.63	.29
Staff-Child Ratio	.057	1	.057	.005	.95
Level of Education	59.12	2	29.56	2.40	.19
2-Way/Int.	36.21	2	18.10	1.47	.31
Residual	61.53	5	12.30	--	--

Four-Year Olds

	SS	DF	MS	F	Sig.
Main Effects	64.23	3	21.41	3.59	.06
Staff-Child Ratio	44.39	1	44.39	7.44	.02
Level of Education	18.99	2	9.50	1.59	.25
2-Way Int.	22.63	2	11.32	1.90	.20
Residual	53.68	9	5.97	--	--

Table 4.11

ANOVA RESULTS ACROSS AGE GROUPS

(Outlier Center Excluded)

PSI	SS	DF	MS	F	Sig.
Main Effects	15.59	3	5.19	1.43	.26
Staff-Child Ratio	9.43	1	9.43	2.59	.12
Level of Education	5.70	2	2.85	.78	.47
2-Way/Int.	7.30	2	3.66	1.00	.39
Residual	72.81	20	3.64	--	--

PPVT					
Main Effects	49.36	3	16.45	1.35	.29
Staff-Child Ratio	21.46	1	21.46	1.76	.20
Level of Education	25.20	2	12.60	1.03	.37
2-Way Int.	15.52	2	7.76	.64	.54
Residual	243.33	9	12.17	--	--

Analyses of variance based on the experimental design also were carried out using the Adult Focus, Child Focus and CDA Checklist data as dependent variables. Although there were scattered significant effects for one or both treatment factors on these variables, there was no strong evidence that either factor had a consistent affect on classroom processes. The significant results are summarized in Table 4.12.

For 53 observation variables, only five (10 percent) showed significant effects ($p<.05$) for one or both factors or for the interaction. When the outlier center was omitted, the pattern of effects changed slightly, with five effects remaining significant and an additional effect emerging. There was no strong pattern of effects within any of the three observation measures; the most consistent pattern was across measures and suggested that either level of education or the interaction between staff/child ratio and level of education affects aide behavior more than it affects teacher behavior, skills or children's behavior. This finding is fairly well substantiated in the replication analyses, which indicated that aides were more affected by the policy variables--especially the classroom composition variables--tnan were teachers or children.

Regression Analyses

In the analyses of variance described above, staff child ratio is categorized as a two-level variable. Observed ratio data obtained during the experimental year suggested that although ratio contrasts were implemented, differences were not as great as expected. The observed data also indicated that staff child ratio was a continuous variable that ranged between 1:4 and 1:10. To incorporate this considerable variation in the experimental analysis, regressions that included observed staff/child ratio and

Table 4.12

SUMMARY OF ANOVA RESULTS FOR AFI, CFI AND CDA VARIABLES

Variable	Source of Variation	F	Sig. of F
All Centers			
Aide Instructs	Education x Staff/Child Ratio	5.84	.01
Aide Management	Education	4.66	.02
Aide Social Interaction	Education x Staff/Child Ratio	7.29	.01
Classroom Management (CDA)	Education x Staff/Child Ratio	4.52	.02
Child Receives Information (Teacher-Directed Activity)	Education	3.29	.05
Omitting Outlier Center			
Aide Instructs	Education x Staff/Child Ratio	4.09	.04
Aide Management	Education	3.81	.05
Aide Social Interaction	Education x Staff/Child Ratio	4.18	.04
Child Receives Information (Teacher-Directed Activity)	1) Education 2) Staff/Child Ratio	4.38 4.89	.03 .04

level of education were computed. To provide additional statistical power, these analyses were combined across age groups using the age-independent change score measure. The results (Table 4.13) confirm the ANOVA results and show that a small staff/child ratio effect exists for the PSI only if all centers are included in the analyses. No significant results are found for the PPVT.

Table 4.13

REGRESSION RESULTS FOR APS DESIGN VARIABLES WITH CHILDREN'S COGNITIVE GAIN SCORES

(N=30)

Dependent Variable	Independent Variable	r	B	SE_B	F	P_F	R^2
PSI							
	Staff/Child Ratio	.36	28.76	13.92	4.27	.05	.13
	Staff/Child Ratio	.36	31.12	14.23	4.78	.02	.16
	Level of Education	.08	.48	.55	.77	.47	
PPVT							
	Level of Education	.25	1.16	.84	1.92	.18	.06
	Level of Education	.25	1.08	.86	1.56	.23	.07
	Staff/Child Ratio	-.14	-12.57	23.15	.30	.74	
PSI							
	Staff/Child Ratio	.23	15.01	12.87	1.36	.25	.05
	Staff/Child Ratio	.23	16.13	13.33	1.46	.25	.06
	Level of Education	.04	.23	.52	.19	.83	
PPVT							
	Staff/Child Ratio	-.29	-34.12	23.32	2.14	.15	.08
	Staff/Child Ratio	-.29	-29.35	23.47	1.56	.23	.14
	Level of Education	.28	1.04	.87	1.41	.26	

CHAPTER FIVE: SUMMARY AND CONCLUSIONS

A classroom-level factorial design that crossed two levels of staff/child ratio and three levels of staff education was implemented in eight of the Atlanta Public Schools day care centers. Although observations of classroom structure indicated that some classes did not maintain assigned ratio levels, the implemented design contained sufficient contrasts (1:5.5 vs. 1:7.5) to merit further analysis. The primary outcome measures used were PSI and PPVT gain scores.

Variance component analyses based on a twofold nested design of classes randomized within centers indicated that no center effect existed for the PSI, but that a significant classroom level effect did exist. A small center effect existed for the PPVT for three-year olds only, but since there was no classroom effect for that dependent measure this effect should not invalidte results from pooled analyses.

Conventional analyses of variance and regression analyses that considered all classes (N=30) showed the presence of a staff/child ratio effect ($p<.05$) for PSI scores of three-year olds. The effect was unchanged when one outlier center was omitted from the analyses. In addition, a staff/child ratio affect was found for PPVT gain scores when the outlier center was excluded. Across age groups, there were no significant effects for the design variables in either the ANOVAs or the regression analyses. When the outlier center was omitted. Thus there is no consistent staff/child ratio effect on cognitive gain scores. Scattered significant ratio or education effects were found for the classroom process variables but the pattern of relationships was not strong. There was no apparent interaction between staff/child ratio and level of staff education except on two aide behaviors. No other effects of either factor were observed for the process measures.

The APS study provides evidence that higher staff/child ratios may be associated with higher PSI gain scores for three-year olds but that level of staff education or for classroom processes has little predictive value either for cogntive changes (PSI, PPVT). This conclusion corroborates independent analyses in other components of the NDCS, which show no consistently strong effect for either staff/child ratio or the two education variables (years of education, highest degree). Given the large classroom effect in the variance components analysis and the small treatment effect in the ANOVA and regression analysis for the PSI, it is likely that a non-experimentally controlled classroom effect contributed to classroom-level differences.

References

Bache, W.L. "Comparing Alternative Measures of Classroom Composition." National Day Care Study Measurements and Methods. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980a.

Connell, D.C. "The Child-Focus Observation Effects Analysis." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Goodrich, N.N. "The Atlanta Public Schools Day Care Experiment." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980a.

Goodrich, N.N. "The Effects of Day Care in Eight Atlanta Public Schools Day Care Çenters." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980b.

Goodrich, N.N. "An Analysis of the CDA Checklist Data." National Day Care Study Measurements and Methods. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980c.

Goodrich, R.L., and Singer, J.D. "Cognitive Change in the NDCS." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Goodson, B.D. "The Adult-Focus Observation Effects Analysis." National Day Care Study Effects Analyses. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Graybill, F., An Introduction to Linear Statistical Models. New York: McGraw-Hill Book Company, Inc., 1961.

The Effects of Day Care in Eight Atlanta Public Schools Day Care Centers

Nancy N. Goodrich

TABLE OF CONTENTS

Page

CHAPTER ONE: INTRODUCTION 435

CHAPTER TWO: VARIABLES AMD METHODOLOGY 439

- Variable Selection 439
- Methodology 441

CHAPTER THREE: ANALYSES OF THE POLICY VARIABLES 443

- Relationships between Policy Variables and Cognitive Gain Scores 443
  - Unit of Analysis 443
  - Construction of Measures of Group Size and Staff/Child Ratio 444
  - Correlation Analyses 444
  - Regression Analyses 444
  - Covariables and Cognitive Gain Scores 448
  - Conclusions 452
- Relationships between Policy Variables and Classroom Process 453
  - Adult-Focus Observations 453
  - Child-Focus Observations 457
  - CDA Checklist Observations 463

CHAPTER FOUR: ANALYSES OF CLASSROOM PROCESS 466

- Relationships between Adult-Focus and Child-Focus Variables 466
- Relationships between CDA and Adult-Focus Variables 470
- Relationships between CDA and Child-Focus Variables 470

CHAPTER FIVE: ANALYSES OF CLASSROOM PROCESS AND CHILDREN'S COGNITIVE GAINS 474

- Relationships between Adult-Focus Variables and Gain Scores 474
- Relationships between Child-Focus Variables and Gain Scores 477
- Relationships between CDA Variables and Gain Scores 479

CHAPTER SIX: CONCLUSIONS 483

References 485

APPENDIX A 487

CHAPTER ONE: INTRODUCTION

The Atlanta Public Schools (APS) study, undertaken as part of the National Day Care Study (NDCS), was a classroom-level randomized experiment that examined two levels of staff/child ratio and three levels of staff education using a factorial design. The study invoiced approximatley 260 tested children and 70 caregivers in eight of the APS day care center. An account of the APS experiment and its conclusions is presented by N. Goodrich (1980). This report expands the study of the APS data to include a number of independent variables which were not experimentally manipulated, but which varied naturally across classrooms.

The APS classroom-level analyses are of interest for several reasons. First, the APS experiment, which investigated the effects of staff/child ratio and level of staff education on a variety of child outcome measures, found that significant differences existed between classrooms but that effects could not be attributed to either of the design factors. Through the replication analyses, it was possible to focus on alternative explanations for these classroom-to-classroom variations. Second, because APS centers serve primarily black children of low socioeconomic status under the same auspices and funding structures, they presented an opportunity to examine policy variables that were not confounded with these major background variables. Finally, Goodrich and Singer (1980) point out that in the 57-center study, the classroom is not the optimal unit of analysis for linking pretest and posttest scores because of frequent classroom reorganization, especially in private day care centers. APS classrooms, however, were relatively stable

---

*These analyses parallel those conducted by Goodrich and Singer (1980), Goodson (1980), Connell (1980) and Singer (1980). All APS replication analyses were performed at the classroom level and examined only those variables included in the APS experimental analysis (see N. Goodrich, 1980).

and permitted the construction of classroom-level cognitive gain scores. Since classroom-level measures of classroom composition and caregiver qualifications may provide better information concerning a child's immediate environment than do center-level measures, the opportunity to perform classroom level analyses appeared very useful.

The principal variables in the study may be classified as follows:

- Policy Variables: Caregiver qualifications (PREVIOUS DAY CARE EXPERIENCE, CENTER EXPERIENCE, LEVEL OF EDUCATION); classroom composition (STAFF/CHILD RATIO, GROUP SIZE, NUMBER OF CAREGIVERS).
- Process Variables: Adult-Focus Instrument (AFI) variables; Child-Focus Instrument (CFI) variables; Child Development Associates (CDA) Checklist variables.
- Child Development Tests: Preschool Inventory and Peabody Picture Vocabulary Test generalized gain scores (PSI GAIN and PPVT GAIN).

The analyses that were conducted include:

- Effects of policy variables on children's and Peabody Picture Vocabulary Test generalized gain scores (PSI GAIN and PPVT GAIN).

The analyses that were conducted include:

- Effects of policy variables on children's generalized PSI and PPVT gain scores.
- Effects of policy variables on classroom process variables.
- Relationships among classroom process variables.
- Effects of classroom process variables on cognitive gain scores.
- Overall relationships among policy, process and child outcome variables.

The major conclusions--which in large part replicate those from the other studies cited--are:

- Importance of Smaller Groups: Children in smaller groups had higher PSI gain scores and were more actively involved with each other and with caregivers. Aides interacted more often with children in these smaller-scale environments.
- Caregiver Interactiveness: Caregivers who interacted more often with children were found in classes where children were more actively involved.
- Differences between Teachers and Aides: Aides appeared to be more strongly influenced by variations in classroom composition than did teachers.
- Caregiver Qualifications: Caregivers with more previous day care experience received higher CDA Checklist ratings. Moreover, classes in which there were more experienced caregivers were characterized by more active participation by children in classroom activities, more attention to children from adults and higher PPVT gains.

In general these findings replicate those of the various three-site studies. It was not possible to determine the effect of previous day care experience in the overall study because of the limited variation in that variable. In addition the APS finding that aides were more influenced than teachers by classroom composition was not replicated in the larger study--a result that probably is due to the unique role that aides played in the APS experiment.

The remainder of this report is organized into five chapters. In Chapter Two, the variables included in the APS analyses and the methodology employed are described. The effects of the policy variables on both classroom process variables and children's generalized cognitive gain scores are presented in Chapter Three. Relationships among process variables are discussed in Chapter Four, and the effects of process variables on cognitive gain scores are presented in Chapter Five. Finally, in Chapter Six the major results of the APS study are summarized. Most of the tables mentioned in the text appear in Appendix A.

CHAPTER TWO: VARIABLES AND METHODOLOGY

Variable Selection

The major variables included in the APS analyses are presented in Figure 2.1. These variables were derived from a common NDCS data base used in both the APS and the three-site studies. However, to take advantage of special characteristics of the APS sample and of the experimental design that was implemented, there were certain differences in variable selection and construction.

First, for most analyses the design variable LEVEL OF EDUCATION was used instead of YEARS OF EDUCATION, HIGHEST DEGREE or SPECIALIZATION. LEVEL OF EDUCATION was created to capture unique educational contrasts that existed only in the APS centers; consequently, it is a more specifically defined measure than are its counterparts in the overall study. A second difference is that averaged measures of GROUP SIZE, STAFF/CHILD RATIO, previous experience, and center experience did not include the September (baseline) data. This computation was revised for the APS sample because the experimental changes required for the study were not effected until October 1976, following the first wave of classroom composition observations. Finally, in the APS study, classroom process analyses focused on variables that indicated what caregivers did and how children responded. In the three-site study, process analyses were supplemented by information relating to more structural processes (such as to whom caregivers directed their attention). Thus the APS analyses include some variables not reported for the various three-site analyses and omit others.

Measures of classroom composition (GROUP SIZE, STAFF/CHILD RATIO, NUMBER OF CAREGIVERS) were available from

Figure 2.1

PRIMARY VARIABLES USED IN APS ANALYSES

Independent Variables

1. Observed Staff/Child Ratio
2. Observed Group Size
3. Observed # Caregivers
4. Level of Education
   1=High School
   2=AAT (Voc/Tech)
   3=Graduate Degree
5. Observed Center Experience (also teacher experience, aide experience)
6. Observed Previous Day Care Experience (Also teacher exoerience, aide experience)

Adult-Focus Variables

1. Commands
2. Direct Questions
3. Responds
4. Instructs
5. Adult Activity
6. Center Activity
7. Comforts
8. Praises
9. Corrects
10. No Response/ Rejects
11. Observes
12. Open Questions

Constructs

13. Management Behavior (1 + 9)
14. Social Interaction (2 + 3 + 4 + 7 + 8)
15. Negative Behavior (from "how" modifiers)
16. Positive Behavior (from "how" modifiers)

CDA Variables

1. Physical Safety
2. Awareness of Safety
3. Sanitation
4. Toys and Equipment
5. Active Play
6. Cognitive/Language
7. Self-Concept
8. Individual Strengths
9. Prosocial Behavior
10. Creative Materials
11. Organizational Skills
12. Class Organization

Constructs

13. Environment (1 + 3)
14. Resources (4 + 10 + 12)
15. Class Management (5 + 11)
16. Child Orientation (2 + 6 + 7 + 8 + 9)
17. CDA Rating (weighted average of all variables)

Child-Focus Variables

1. Monitors Environment
2. Wanders
3. Gives Opinions
4. Gives Orders
5. Receives Orders
6. Receives Information
7. Receives Generally

Constructs

8. Participates in Open Activity (participates in open activity alone and participates in open activity in group)
9. Participates in Structured Activity (participates in structured activity alone and participates in group)

Child-Outcome Variables

1. PSI Gain Score
2. PPVT Gain Score

three sources: observation data averaged from several time points; counts made during adult-focus observations (April-May 1977); and counts made during child-focus observations (April-May 1977). Relationships among data from these three sources have been examined for both the APS and the larger studies (Bache, 1980). Because low correlations were found between the adult-focus counts and the other measures,* and because caregivers' behaviors are likely to be influenced by the classroom organization at the time of the observation, the adult-focus counts were used in both the AFI and the CDA analyses; the averaged observation data were used in the remaining analyses.

The data measuring classroom processes (AFI, CFI, CDA) and children's cognitive development (PSI, PPVT) were collected in October 1976 (T3) and May 1977 (T4) by observers and testers from International. Information about the policy variables was collected on an on-going basis by center secretaries provided by Abt Associates. Abt's staff were responsible for defining, editing, and analyzing all of the variables reported in this paper.

Methodology

Relationships among policy, process and outcome variables were investigated primarily through use of multiple regression techniques. Correlations were also computed for preliminary investigations of relationships among variables. In addition, for the process-outcome analysis, partial correlations were employed to adjust for effects of relevant policy variables.

---

*Adult-focus staff and child counts were slightly lower than the other observation counts, suggesting that different criteria may have been used to define "group" in the adult-focus observations.

The focus of the regression analyses was to develop a simple predictive model without introducing severe multicollinearity. Forward stepwise regression using the algorithm in the SPSS software package was the principal investigative tool. Regressions to predict cognitive gain scores were weighted by the number of tested children in the classroom.

As variables were entered into the regression model, changes both in the regression coefficients (B's) and the percent of variance accounted for (R^2) were examined. In addition, the significance of the model was tested using the F-ratio. Multicollinearity was presumed to be present if any of the B coefficients fluctuated by more than one standard deviation when a new variable was introduced. Further, each individual variable was required to be significant at approximately the .05 level. For some borderline cases, the decision to include or exclude a variable was based on intuition and/or the extent to which the resulting model paralleled models for other dependent variables. The models that emerge do not always achieve the highest possible R^2 (given the estimated reliabilities of the measures), but regression coefficients of the models are readily interpretable as "effects" (although causal inferences cannot be made from associational data).

CHAPTER THREE: ANALYSES OF THE POLICY VARIABLES

Relationships Between Policy Variables and Cognitive Gain Scores

Generalized gain scores for two cognitive tests--the Preschool Inventory (PSI) and the Peabody Picture Vocabulary Test (PPVT)--were used in the APS analyses. Although data were available for a gross motor scale based on the Denver Developmental Screening Test and for the Pupil Observation Checklist (POCL), psychometric analyses (Bache, 1980) indicated that these change scores were not sufficiently reliable to be analyzed. The generalized gain scores used in the APS study are identical to those used in the 57-center cognitive effects analyses (Goodrich and Singer, 1980). The algebraic formulas used to compute gain scores are:

$$\text{PSI GAIN} = \text{T4PSI} - (.88 \times \text{T3PSI})$$

and

$$\text{PPVT GAIN} = \text{T4PPVT} - (.88 \times \text{T3PPVT}).$$

Unit of Analysis

Two units of analysis--child and classroom--were considered initially. Class-level regressions and correlations were weighted by the number of tested children per class. When independent variables are defined only at class level, child-level and weighted class-level regression coefficients are mathematically identical (see Goodrich and Singer, 1980). However, hypothesis testing is different in the two cases and is not valid at child level if (as is usually true) the residuals are correlated within-class. Regressions run at class and child level (Tables A.1 and A.2, Appendix A) provided empirical verification of these theoretical results. Regression analyses were therefore conducted at the classroom level.

Construction of Measures of Group Size and Staff/Child Ratio

In a second analysis, the appropriate construction of group size and staff/child ratio measures was investigated (see also Bache, 1980). Regression models were developed for both the untransformed classroom composition variables and for the logarithms of these variables, using generalized gain scores as dependent measures. The results (Tables A.3 and A.4, Appendix A) showed that the unlogged classroom composition variables were slightly stronger predictors than were their logged counterparts. Consequently, basic regression models were developed for the untransformed measures.

Correlation Analyses

Correlations of the policy variables with each other and with PSI and PPVT gain scores were computed at the classroom level weighting by the number of children tested (Tables 3.1 and 3.2). GROUP SIZE and STAFF/CHILD RATIO were both significantly correlated ($p<.05$) with PSI GAIN; PREVIOUS DAY CARE EXPERIENCE was the only policy variable significantly correlated with PPVT GAIN. NUMBER OF CAREGIVERS was highly correlated with both GROUP SIZE and STAFF/CHILD RATIO; however, STAFF/CHILD RATIO and GROUP SIZE were relatively independent ($r=.17$, $p>.15$). Both PREVIOUS DAY CARE EXPERIENCE and CENTER EXPERIENCE were correlated with GROUP SIZE, creating potential problems in interpretation.

Regression Analyses

To investigate the relationship between policy variables and generalized gain scores, regression equations were estimated using all possible combinations (one to six variables) of the following policy variables:

Table 3.1

PEARSON CORRELATIONS FOR POLICY VARIABLE AND PSI GAIN SCORES

Independent Variable	PSI	Years of Education	Staff/Child Ratio	Group Size	Center Exp.	Previous Day Care Exp.	Specialization	Level of Education	
Years of Education		1							
Staff/Child Ratio	.21**	.11*	1						
Group Size	-.35**	-.10*	-.16**	1					CHILD LEVEL
Center Exp.			-.10*	.29**	1				N=265
Prev. Day Care Exp.			-.26**	-.33**		1			
Specialization	.08		-.23**		.26**	.12*	1		
Level of Education		.60**	-.12*	-.12*	.26**	.08	.09		
Number of Caregivers	-.10*	-.10*	.50**	.68**	.18**	-.38**			
Years of Education		1							
Staff/Child Ratio	.36*		1						
Group Size	-.62**			1					
Center Exp.				.28	1				WEIGHTED AT
Prev. Day Care Exp.			-.26	-.33*		1			CLASS LEVEL
Specialization					.26		1		N=30
Level of Education		.50**							
Number of Caregivers			.49**	.68**		-.37*			

---

*$p<.05$

**$p<.01$

Correlations are reported only for $p<.15$.

Table 3.2

PEARSON CORRELATIONS FOR POLICY VARIABLE AND PPVT GAIN SCORES

CHILD LEVEL
N=259

Independent Variable	PPVT	Years of Education	Staff/Child Ratio	Group Size	Center Exp.	Previous Day Care Exp.	Specialization	Level of Education
Years of Education		1						
Staff/Child Ratio		.11	1					
Group Size	-.09	-.13*	-.15*	1				
Center Exp.			-.12*	.31**	1			
Prev. Day Care Exp.	-.21**		-.26**	-.32**		1		
Specialization	.11		-.25**		.25**	.16*	1	
Level of Education		.60**	-.13*	-.10	.26**		.11	1
Number of Caregivers		-.12*	.50**	.69**	.19**	-.39**		-.10

WEIGHTED AT
CLASS LEVEL
N=30

Independent Variable	PPVT	Years of Education	Staff/Child Ratio	Group Size	Center Exp.	Previous Day Care Exp.	Specialization	Level of Education
Years of Education		1						
Staff/Child Ratio			1					
Group Size	-.23			1				
Center Exp.				.31*	1			
Prev. Day Care Exp.	-.51**		-.26	-.33*		1		
Specialization	.26		-.25		.25		1	
Level of Education	.25	.50**	-.19					1
Number of Caregivers			.49**	.68**	-.19	-.38*		

---

*p<.05

**p<.01

Correlations are reported only for p<.15.

- GROUP SIZE;
- STAFF/CHILD RATIO;
- YEARS OF EDUCATION;
- SPECIALIZATION;
- CENTER EXPERIENCE;
- PREVIOUS DAY CARE EXPERIENCE; and
- LEVEL OF EDUCATION

Another policy variable--NUMBER OF CAREGIVERS--also was considered in preliminary regressions. This variable, however, was not as good a predictor as either GROUP SIZE or STAFF/CHILD RATIO, and because any two of these three variables specify the third variable (i.e., STAFF/CHILD RATIO = GROUP SIZE/NUMBER OF CAREGIVERS), all three could not be entered simultaneously.* The investigation was therefore pursued using only STAFF/CHILD RATIO and GROUP SIZE.

Two different regression models emerged from these analyses (see Tables A.1 and A.2, Appendix A). For the PSI, the model included the combination of GROUP SIZE, STAFF/CHILD RATIO and SPECIALIZATION ($R^2 = .51$) whereas the PPVT model utilized only PREVIOUS DAY CARE EXPERIENCE ($R^2 = .26$). Adding LEVEL OF EDUCATION and CENTER EXPERIENCE to the PPVT model increased the R^2 to .34. However, these variables were not included in the model because neither was a significant predictor either alone or with PREVIOUS DAY CARE EXPERIENCE.

To determine whether major regression results could be attributed to outlier classrooms, robust (biweighted) regressions also were performed. The biweighted regression iteratively down-weights the influence of classes that are further from an ordinary least squares regression surface. Although several outlier classes were identified, omitting these classes from the analysis did not significantly change the regression results. Classroom composition measures, especially GROUP SIZE, remained significant predictors of PSI GAIN, and PREVIOUS DAY CARE EXPERIENCE remained a significant predictor of PPVT GAIN.

---

*If all three measures were included, there would be a linear combination among the independent variables; with such a linear combination, it would be impossible to estimate regression models.

Both the correlation and regression results indicate that children's PSI and PPVT gain scores are associated with different policy variables. GROUP SIZE is strongly associated with PSI GAIN but not with PPVT GAIN. On the other hand, PREVIOUS DAY CARE EXPERIENCE is positively and significantly related to PPVT GAIN but not to PSI GAIN. It is perhaps not surprising that two distinct regression models emerged for predicting PSI and PPVT gain scores. Although PSI and PPVT scores are highly correlated in the APS sample ($r_{T3} = .62$; $r_{T4} = .59$), their gain scores are not significantly correlated ($r_{GAIN} = .06$). Thus, there is evidence that the two tests measure different aspects of children's cognitive development.

Covariables and Cognitive Gain Scores

The effects of nine child-specific covariables on PSI and PPVT generalized gain scores also were investigated in the APS analyses. These covariables were:

- NUMBER OF ADULTS (in the household);
- NUMBER OF SIBLINGS;
- NUMBER OF SIBLINGS UNDER 12 YEARS;
- AGE OF CLOSEST SIBLING;

- FAMILY INCOME (categorical variable);
- MOTHER'S EDUCATION;
- AGE;
- RACE; and
- SEX.

The influence of these covariables on children's cognitive development was examined at both child and classroom levels. At the child level, covariables and fall test scores were used to predict spring test scores. The results (Table 3.3) indicate that for both the PSI and the PPVT the fall test score was a much stronger predictor of the spring score than were any of the covariables. Although several of the covariables were statistically significant predictors (NUMBER OF ADULTS and AGE for the PSI score; AGE and SEX for the PPVT score), their overall significance at the child level was not of practical concern. When both the covariables and the policy variables were entered into regressions predicting generalized gain scores at the child level, NUMBER OF ADULTS was a barely significant predictor of PSI GAIN ($p<.05$); there were no significant predictors for PPVT GAIN.

Covariables also were used to predict age-adjusted gain scores at the classroom level. These aggregated variables thus became measures of the learning environment rather than of an individual child. It is not surprising, therefore, that different regression results were obtained at the classroom level (Table 3.4). AGE OF CLOSEST SIBLING was a significant predictor of PSI GAIN, and MOTHER'S EDUCATION was a significant predictor of PPVT GAIN. The significance of these covariables was probably due to the fact that they were highly correlated with the policy variables included in the regression models for predicting gain scores (for GROUP SIZE/AGE OF CLOSET SIBLING $r= -.48$; for STAFF/CHILD RATIO/AGE OF CLOSEST SIBLING $r=.32$; for PREVIOUS DAY CARE EXPERIENCE/MOTHER'S EDUCATION $r=.36$).

Table 3.3

COVARIABLE REGRESSION RESULTS: CHILD LEVEL

Dependent Variable	Independent Variable	B	SE_B	F	R^2
Spring PSI	Fall PSI	.77	.05	199.93**	.45
	Fall PSI	.78	.05	210.08**	.46
	Number of Adults	1.19	.40	8.70**	
	Fall PSI	.71	.06	132.32**	.48
	Number of Adults	1.15	.40	8.31**	
	Age	1.29	.57	5.14**	
Spring PPVT	Fall PPVT	.76	.05	194.43**	.44
	Fall PPVT	.68	.06	119.98**	.45
	Age	3.32	1.16	8.12**	
	Fall PPVT	.67	.06	119.98**	.45
	Age	3.55	1.17	9.23**	
	Sex	1.83	1.06	3.01*	
PSI Gain Score	Number of Adults	.91	.43	4.38*	.02
	Number of Adults	.95	.43	4.88*	.03
	Age of Next Sibling	.11	.07	2.57	
PPVT Gain Score	Number of Adults	.89	.91	.94	.004

---

*p<.05
**p<.01

Table 3.4

COVARIABLE REGRESSION RESULTS: CLASSROOM LEVEL

Dependent Variable	Independent Variable	B	SE_B	F	R^2
PSI Gain Score	Age of Next Sibling	.53	.22	5.75*	.17
	Age of Next Sibling	.52	.21	6.05*	.26
	Number of Adults	2.34	1.31	3.18	
	Group Size	.35	.08	18.83**	.56
	Staff/Child Ratio	15.17	11.69	1.68	
	Center Experience	.48	.26	3.34*	
	Previous Day Care Experience	-.75	.48	2.48	
	Number of Adults	1.24	1.18	1.11	
PPVT Gain Score	Mother's Education	6.87	2.10	10.72**	.28
	Mother's Education	7.27	2.11	11.81**	.31
	Number of Siblings <12 years	1.43	1.24	1.32	
	Previous Day Care Experience	2.31	.74	9.72**	.26
	Mother's Education	5.12	2.11	5.87**	
	Previous Day Care Experience	1.65	.73	5.02*	

---

*$p<.05$
**$p<.01$

When the covariables were entered into regressions with the policy variables, the overall results did not change. The stepwise regression procedure did not select any covariables for the PSI model, and although MOTHER'S EDUCATION was a better predictor of PPVT GAIN than was PREVIOUS DAY CARE EXPERIENCE, the experience variable still remained significant after mother's education was considered. Thus, the major results do not appear to be threatened by covariable effects.

Conclusions

The APS results indicate that children's PSI gain scores are positively associated with small-scale class composition, especially small group sizes, and that PPVT gain scores are more related to caregiver backgrounds. In particular, PPVT results suggest that caregivers with more previous day care experience are found in those classes where there is greater improvement in children's vocabulary skills as measured by the PPVT.* Several background variables (AGE OF CLOSEST SIBLING, MOTHER'S EDUCATION, PREVIOUS DAY CARE EXPERIENCE) are also predictive of cognitive gains. However, the principal conclusions remain stable when these covariables are controlled for in regression models.

---

*It should be noted that PSI results tend to indicate that even though the two experience variables are not significant predictors of PSI gains, less previous day care experience and more center experience are related to higher gain scores. This finding does not appear to be a result of outlier values for either of the two independent variables; in addition, the unweighted distribution of variables is similar for both gain scores. Thus it seems probable that the anomalous relationship is due more to inadvertent confounding of the experience variables with GROUP SIZE than to differences in cognitive gains as measured by the PSI and the PPVT.

Relationships between Policy Variables and Classroom Process

Adult-Focus Observations

Data obtained from the Adult-Focus Instrument (AFI) describe caregiver activity in the classroom using "sentences" identifying an actor (who), the object of the action (to whom), the nature of the action (what), and the style of the action (how). For analytic purposes, the information in these sentences was reduced to the percentage of observed time that the focal caregiver was involved in each activity type (Goodson, 1980). The Adult Focus Instrument includes a Physical Environment Inventory, which describes space, materials and equipment in the classroom; a Classroom Snapshot, which describes general activity patterns at a point in time; and a Five-Minute Interaction record, which describes the behavior of a particular caregiver in detail. The Five-Minute Interaction data are of primary interest here.

Because of the unique educational comparisons created in the APS study design, the analyses focused primarily on a selected set of "what" codes to investigate differences in caregiver activities that might exist for different levels of education and job roles. Factor analyses, correlations and conceptual considerations were used to determine a set of six variables for the investigation:

- INSTRUCTS;
- ADULT ACTIVITY;
- CENTER ACTIVITY;
- OBSERVES;
- MANAGEMENT BEHAVIOR (defined as COMMANDS, CORRECTS); and
- SOCIAL INTERACTION (defined as DIRECT QUESTIONS, RESPONDS, INSTRUCTS, COMFORTS, PRAISES).

The MANAGEMENT BEHAVIOR and SOCIAL INTERACTION constructs were defined as in the 57-center adult-focus analyses (Goodson, 1980). APS factor analyses suggested, however, that INSTRUCTS was less related to social interaction than were the other codes included in this construct (DIRECT QUESTIONS, RESPONSES, COMFORTS, PRAISES). The APS analysis therefore also examined INSTRUCTS independently.

Values for classroom composition variables used in adult-focus analyses were derived from the Classroom Snapshot immediately preceding the Five-Minute Interaction. GROUP SIZE, NUMBER OF CAREGIVERS and STAFF/CHILD RATIO were computed by averaging staff and child counts taken before each five-minute observation. PREVIOUS DAY CARE EXPERIENCE, CENTER EXPERIENCE and LEVEL OF EDUCATION data are for the focus caregiver. Teacher- and aide-focus data were examined separately in all analyses. Aide-focus observations were available for only 22 of the 23 classes with aides; teacher-focus observations were available for all 30 lead teachers.

Regression models were developed using policy variables as predictors for each of the six adult-focus constructs. The results are shown in Table 3.5. For teachers, the regressions for the adult-focus variables suggest that a level-of-education effect exists for both CENTER ACTIVITY and OBSERVES. Since all teachers with a high level of education (M.A.) also served as center directors, it seems probable that the education effect was due more to job requirements than to education. To investigate this rival hypothesis, regressions were re-estimated excluding center directors; as expected, the level of education effect disappeared. Without the directors, the only adult-focus variable that was significantly predicted by any of the policy variables was CENTER ACTIVITY, for which a combination of STAFF/CHILD RATIO and TOTAL DAY

Table 3.5

SUMMARY OF POLICY VARIABLE/ADULT FOCUS REGRESSION RESULTS

	Predictor	All Teachers Sign	R^2	Predictor	AAT, HS Teachers Sign	R^2		Predictor	Aides Sign	R^2
Instructs								Group Size**	(-)	.32
Adult Activity								S/C Ratio*	(-)	.22
Center Activity	Ed. Level S/C Ratio*	(+) (+)	.17 .25	S/C Ratio* Center Exp. Prev. D.C. Exp.	(+) (+) (+)	.23 .41[1]				
Observes	Ed. Level Group Size	(-) (-)	.17 .26					Group Size**	(+)	.44
Management Behavior							A)	D.C. Exp.* Group Size: S/C Ratio*	(-) (-) (-)	.25 .41 .50
							B)	D.C. Exp.* Ed. Level* # Caregivers**	(-) (+) (-)	.25 .36 .48
							C)	Group Size* Center Exp.* S/C Ratio*	(-) (-) (-)	.23 .38 .57
Social Interaction								Group Size**	(-)	.57

---

[1] Previous day care experience and center experience are correlated -.47 in the sample. Consequently, their effects cannot be separated. Their sum also is a significant predictor together with staff/child ratio (R^2 = .407).

CARE EXPERIENCE (CENTER EXPERIENCE plus PREVIOUS DAY CARE EXPERIENCE)* accounted for 41 percent of the variance.

The classroom composition variables were much stronger predictors of aide behavior than of teacher behavior. GROUP SIZE was the only significant (p<.01) predictor of three of the variables--INSTRUCTS, OBSERVES and SOCIAL INTERACTION--and was the best single predictor of MANAGEMENT BEHAVIOR. STAFF/CHILD RATIO was the only significant (p<.05) predictor of ADULT ACTIVITY and, together with CENTER EXPERIENCE and GROUP SIZE, was significant in predicting MANAGEMENT BEHAVIOR. Thus, there is evidence that aides in the APS study were especially affected by classroom composition. Aides tended to be more active in smaller classes with higher ratios. As classes became larger and ratios decreased, aides became passive and tended to observe more often or to engage in adult-related activity. Although the classroom composition variables were not significant predictors of teacher activities, in general the same pattern of behavior was observed: teachers were more active in smaller-scale environments. (Details of the regression analyses are shown in Tables A.5 and A.6, Appendix A).

For all adult-focus constructs except aides' MANAGEMENT BEHAVIOR, there was at most one plausible predictive model. It is evident that the model for MANAGEMENT BEHAVIOR for aides should include some combination of the composition variables and CENTER EXPERIENCE (or TOTAL DAY CARE EXPERIENCE); the best model may also include LEVEL OF EDUCATION. Although a single model cannot be identified due to the similar regression results for several independent variables, it is clear that aides' management behavior is affected by both classroom composition and caregiver qualifications; moreover, it is the only aide-focus variable for which caregiver background information is a significant predictor.

---

*These two variables are correlated -.47 in the sample. Because of the multicollinearity, they cannot be separated in this particular regression.

Overall, the regression analyses gave little indication that caregiver behaviors are associated with the caregiver's level of education. However, the relatively small sample size (N=30 for teachers, N=22 for aides) may not have provided sufficient power to detect small differences that might exist. Since the APS design presented a unique opportunity to study the interaction between the level of education and job roles (teacher or aide) of caregivers, the adult-focus data were examined extensively for interesting patterns of behavior by education within job. Analyses included graphic techniques such as stem-and-leaf displays and schematic plots, t-tests for differences in means, and correlations by level of education and job role.

The results, which are shown in Tables A.7 to A.11 in Appendix A, show that although there was some indication that caregivers with more education were more involved in classroom activities, the evidence did not support a level of education effect. The most consistent pattern to emerge was that caregivers whose job roles were changed for the experimental year were less vulnerable to the effects of a large-scale classroom environment. This may be a genuine result of a change in routine, or it may be an indication that the experimental situation affected caregiver behaviors (Hawthorne effect).

Child-Focus Observations

Child-Focus Instrument (CFI) observations were obtained for target children in all 30 APS classrooms. The CFI was administered on three different mornings during a

CFI was administered on three different mornings during a 20-minute observation period divided into 100 12-second coding intervals. Observers were provided with timers that clicked every 12 seconds and were instructed to record what was happening to a selected focus child at the time of the click. For each coding interval, a record was made of the specific activity in which the child was engaged and the object of the child's attention. As for the AFI, scores were computed as proportions of time spent in each specific activity. Children were observed during teacher-directed activities and free-play situations. Data for each activity type were analyzed both separately and combined across the two situations. These combined variables are believed to reflect overall behavior patterns during the day-care day.

Nine CFI variables were selected for investigation in the Atlanta Public Schools study:

- MONITORS ENVIRONMENT (child looks, watches; does not include listening);
- WANDERS (child has no apparent purpose in his movement about the center);
- GIVES OPINIONS (child initiates statements about his own likes, dislikes or preferences);
- GIVES ORDERS (child issues command to another child or caregiver);
- RECEIVES ORDERS (child receives commands with which compliance is expected);
- RECEIVES INFORMATION OR HELP (child receives instruction, materials or assistance related to his task in the solution of his problem);
- RECEIVES GENERAL COMMENTS (child is asked for information or receives comments of a general nature);
- PARTICIPATES IN OPEN ACTIVITY (child is involved in an activity with no defined goal); and
- PARTICIPATES IN STRUCTURED ACTIVITY (child is involved in an activity that has a goal, clear guidelines for carrying out the task, and a defined beginning and end).

Preliminary analyses of the CFI data (Connell, 1980) indicated that the observations were not reliable measures of individual child behaviors, but that they were accurate descriptors of classroom process. Therefore, in the APS, as in the 57-center study, CFI data were analyzed at the classroom level.

To investigate relationships among the policy variables and the CFI data, correlations were computed for several alternative measures of the policy variables. Results are presented in Table 3.6. Although there are isolated instances of inconsistency across variables, the overall pattern of correlations was strong and consistent. GROUP SIZE was highly correlated with children's behavior, as were lead teacher's PREVIOUS DAY CARE EXPERIENCE and lead teacher's CENTER EXPERIENCE. NUMBER OF CAREGIVERS also was correlated with children's behaviors, with approximately the same pattern of correlations as was evidenced by GROUP SIZE, although the correlations were not as strong. STAFF/CHILD RATIO, TOTAL DAY CARE EXPERIENCE and LEVEL OF EDUCATION, however, did not show consistently strong relationships with the CFI measures.

PREVIOUS DAY CARE EXPERIENCE and CENTER EXPERIENCE both showed stronger correlations for teachers than for aides; otherwise the two variables behaved quite differently. PREVIOUS DAY CARE EXPERIENCE seemed to be more important during teacher-directed activities, while the CENTER EXPERIENCE correlations were higher for free-play situations. In addition, the signs of the correlations were usually opposite for the two experience variables. The pattern suggests that lead teachers with varied experience in day care are found in classrooms in which children are more involved, especially during teacher-directed activities, and conversely, that lead teachers who have worked in only one setting over a period of time are found in classes where children are less involved, particularly during free play.

Table 3.6

PEARSON CORRELATIONS FOR CHILD FOCUS AND POLICY VARIABLES

(n=28)

	Group Size				Number of Caregivers			Staff/Child Ratio				Level of Education		Day Care Experience			Previous Day Care Experience			Center Experience		
	Child Focus	Teacher Focus	Adult Aides	Four Point	Teacher Focus	Adult Aides	Four Point	Child Focus	Teacher Focus	Adult Aides	Four Point	Teacher	Aide	Class	Teacher	Aide	Teacher	Aide	Observed	Teacher	Aide	Observed
Monitors Envir.																						
Comb.	.30*	.38*	.49**	.47**			.30*		-.25		-.21		-.27				-.32*	-.26	-.35*	.28	.38	
T. Dir.	.19	.41**	.52**	.46**			.48**		-.23	-.34							-.32*	-.31	-.38*	.26		
Free Play				.31*				-.28	-.22		-.25		-.38*				-.28		-.30*		.30	
Wonders																						
Comb.	.63**	.45**	.29	.34*			.48**		-.19		.22	-.28					-.47**		-.52*	.31	.25	.35*
T. Dir.	.60**	.39*	.47**	.46**			.48**		-.30*	-.31	.19						-.37*		-.39*	.29		.26
Free Play	.30*	.25			.23							-.39*					-.23		-.28			.24
Gives Opinions																						
Comb.	-.33*	-.49**	-.44*	-.38*	-.41**	-.48*	-.35*	.21					.27				.38*		.30	-.27		-.37*
T. Dir.	-.33*	-.19	-.41*	-.32		-.47**		.21					.27				.21	-.27				
Free Play	-.30*	-.65**		-.36*	-.41**		-.40**	.22						-.28	-.23				.22	-.27		-.38*
Gives Orders																						
Comb.				.33*			.22		-.24	-.42*	-.26	.22		.31*	.25					.20		.34*
T. Dir.				.32*				-.30	-.23	-.35*		.38*							-.21			
Free Play								-.22		-.39*	-.21			.26							.33	.29
Receives Orders																						
Comb.	-.45**	-.26	-.41*	-.37*		-.40*	-.29					-.25			-.20	-.32	.45**		.28	-.46**	-.32	-.38*
T. Dir.	-.12					-.30				-.25		-.34*				-.27	.48*		.33*	-.20	-.32	-.26
Free Play	-.42**	-.36*	-.57**	-.48**	-.19	-.40*	-.33*							-.20	-.28	-.24	.38			-.41**		-.35*
Receives Info.																						
Comb.	-.67**	-.37*	-.65**	-.72**			-.49**	.22	.29	.48*							.43**		.40**	-.36		-.19
T. Dir.	-.39**	-.28	-.37*	-.52**		.28	-.31*	.23	.27	.49**				.29			.43**		.53**			
Free Play	-.33*	-.20	-.38*	-.32*		-.28		.27				.25		-.20	-.22	-.27	.23			-.34*		-.29
Receives Gen'l.																						
Comb.		-.29			-.51**	-.52**							.30	-.27		-.25	.20					-.46**
T. Dir.					-.28	-.47**							.26	-.29		-.23			-.19			-.31*
Free Play		-.33*			-.49**	-.41*	-.24				-.19		.26	-.21			.19		.22	-.20		-.41**
Open Activity																						
Comb.	.56**	.33*	.54**	.45**		.40*	.25	-.22					-.46*	.50**	.31*	.45**	-.22			.41**	.42*	.55**
T. Dir.	.22																	-.34	-.35*			
Free Play	.49*	.37*	.39*	.21	.30	.40*		-.31*					-.32	.55**	.35*	.47**				.37	.46*	.55**
Structured Act.																						
Comb.	-.37*	-.22	-.43*	-.38*			-.23		.35*				.34	-.22	-.28				.19	-.28	-.37*	-.25
T. Dir.	-.31*		-.47**	-.41*	.28		-.27		.36*	.31					-.21			.20	.21	-.20	-.51*	-.38*
Free Play	-.30*							.22	.19			.23	.30	-.65**	-.32*	-.34				-.23		

For aide data, n=21

* p<.05
**p<.01

Correlations are reported only for p<.15

Regression models also were developed to predict each of the nine CFI variables. Correlation analyses had indicated that the observed policy variables generally were at least as highly correlated with children's behaviors as the child-focus or individual caregiver variables. Since the Child-Focus Instrument did not record which caregivers were present during the actual child-focus observations, the observed background variables were felt to provide a more accurate representation of caregiver qualifications for the observation period. Therefore, the observed classroom composition and staff background variables were used for regression model development.

The results of the regression analyses are summarized in Table 3.7; details of the analyses are provided in Tables A.12 to A.14, Appendix A. Separate regressions were run for GROUP SIZE and NUMBER OF CAREGIVERS since these two variables were highly correlated in the sample (r=.70). Other variables included in the regressions were LEVEL OF EDUCATION, STAFF/CHILD RATIO, PREVIOUS DAY CARE EXPERIENCE and CENTER EXPERIENCE.

Although no single model consistently emerged for all the child-focus variables, several patterns are evident. GROUP SIZE was the most consistently significant predictor: children in smaller groups received information and orders more often, gave more opinions and participated in structured activities more frequently. Children in larger groups, on the other hand, were more likely to monitor the environment and wander. Caregivers with more previous day care experience were found in classes in which children were more involved while caregivers with more center experience had children who were less involved. These results paralleled the findings suggested by the correlation analyses.

484

Table 3.7

SUMMARY OF POLICY VARIABLES: CHILD FOCUS REGRESSION RESULTS

	Teacher-Directed			Free Play			Combination		
	Variables	Sign	R^2	Predictor	Sign	R^2	Predictor	Sign	R^2
Monitors Environment	Group Size* [# Caregivers]	(+) (+)	.21 .18	(Group Size)	(+)	.11	Group Size**	(+)	.24
Wanders	Group Size* S/C Ratio* [# Caregivers	(+) (+) (+)	.21 .30 .16]	(Center Exp.)	(+)	.10	Prev. D.C. Exp. (Center Exp.)	(-) (+)	.26 .35
Gives Opinions	(Grp. Size)	(-)	.10	# Caregivers*	(-)	.16	(Group Size)	(-)	.15
Gives Orders	(Group Size)	(+)	.10	(Center Exp.)	(+)	.079	(Center Exp.)	(+)	.11
Receives Orders	(Prev. D.C. Exp.)	(+)	.12	Group Size**	(-)	.22	Center Exp.	(-)	.14
Receives Info., Help	Prev. D.C. Exp.** Group Size	(+) (-)	.28 .42	(Group Size)	(-)	.10	Group Size** [# Caregivers**	(-) (-)	.52 .25]
Receives General	(Center Exp.)	(-)	.098	Center Exp.*	(-)	.17	Center Exp.*	(-)	.20
Part. in Open Act.	(Prev. D.C. Exp.)	(-)	.12	Center Exp.**	(+)	.30	Center Exp.**	(+)	.30
Participate in Structured Act.	Group Size* [(# Caregivers)	(-) (-)	.17 .078]	Center Exp.*	(-)	.15	(Group Size)	(-)	.15

---

*p<.05
**p<.01
( ) not significant, but entered first.

CDA Checklist Observations

Caregivers' skills and behaviors were also measured by means of a checklist developed for the NDCS based on the Child Development Associates (CDA) credentialing system (N. Goodrich, 1980). Items on the NDCS checklist were culled from a number of actual CDA observations recorded as part of the credentialing process; the items were organized to reflect the functional areas specified by the CDA. Factor analyses were performed within areas to identify one or two distinct factors to describe caregivers skills. These factors subsequently were factor-analyzed to identify four major constructs--child orientation, classroom management, resources, and physical environment. In addition, an overall CDA rating scale was developed from the original factors.

The CDA Checklist was completed by observers following the AFI obseration period. Only one Checklist was recorded per day, ostensibly for the lead teacher although two caregivers--a teacher and an aide--might have been observed with the AFI. Thus, the Checklist data probably reflect the skills and behavior of more than one caregiver. Given this situation, the CDA instrument must be viewed as a measure of classroom process and not of individual caregivers.

Correlations with the policy variables were computed for the five constructs and the twelve factors defined for the CDA analyses (Table 3.8). Because of the high correlations among the CDA factors and the difficulties in interpreting figures for twelve different factors, results are discussed for the five constructs; data for the twelve factors are presented only to provide supporting evidence.

The pattern of correlations suggests that smaller classes, higher ratios and more previous day care experience

Table 3.8

PEARSON CORRELATIONS FOR POLICY AND CDA VARIABLES

(N=30)

	CDA Rating	Child Orientation	Encourages Active Play	Encourages Cog. Long. Dev.	Encourages Good Self-Concept	Encourages Self Help	Encourages Social Behavior	Classroom Management	Encourages Safety	Manages Class Act. Well	Resources	Provides Gross Motor Toys	Provides Creative Play Mat.	Arranges Classroom Well	Physical Environment	Maintains Safe Classroom	Maintains Sanitary Classroom
Group Size	-.35*	-.40*	-.19	-.56**	-.44**	.19	-	-.47**	-.31*	-.47**	-	-	-	-.22	-	-	-.32*
Number of Caregivers	.21	-	-	-	-	.54**	-	-	-	-	.37*	-	.56**	.21	.30	.36*	-
Staff/Child Ratio	.43**	.39*	-	.46**	.33*	.26			-	.25	.40*	-	.49**	.36*	.45**	.30	.39*
Center Exp.	-	-	.33*	-	-	.31*	-	-	.20	-	-	-	-	-	-.22	-	-.27
Previous Day Care Experience	.52**	.48**	.62**	.34*	.46**	-	.54**	.51**	.19	.58**	.41*	.50**	-	.43**	-	-	.37*
Level of Education			-	.37*	.20	-	.25	.36*	.32*	.33*	-	-	-	-	-	-	-

---

*p<.05
**p<.01
Correlations are reported only for p <.15.

for caregivers are associated with higher CDA ratings. No significant correlations were found for CENTER EXPERIENCE with any of the factors or constructs. NUMBER OF CAREGIVERS showed a few high correlations but was not as strongly associated with the CDA measures as were the other classroom composition variables. The overall pattern is consistent with the AFI results, which indicated that caregivers were more active in smaller scale environments.

Regression models for each factor and construct are displayed in Table A.15, Appendix A. Both PREVIOUS DAY CARE EXPERIENCE and STAFF/CHILD RATIO were good predictors of caregivers' behavior and were the only policy variables included in the regression model. STAFF/CHILD RATIO was more closely associated with the classroom environment (PHYSICAL ENVIRONMENT; RESOURCES) while PREVIOUS DAY CARE EXPERIENCE was the strongest predictor of caregiver activity (CLASSROOM MANAGEMENT; CHILD ORIENTATION). Despite significant correlations with many of the CDA variables, GROUP SIZE did not emerge in the regressions as a good predictor of caregivers' skills and behaviors. Since the teacher-focus measures of GROUP SIZE and STAFF/CHILD RATIO had a correlation of -.44, the effects of STAFF/CHILD RATIO simply outweighed the effects of GROUP SIZE in the forward stepwise regression used in the analysis.

CHAPTER FOUR: ANALYSES OF CLASSROOM PROCESS

Relationships between Adult-Focus and Child-Focus Variables

Although the AFI and CFI observations were made during the same time period (April-May 1977), they were not conducted simultaneously. Thus, there is no direct record of how children and caregivers interacted with each other. Since both the AFI and CFI data represent averages of behaviors over several days, it is reasonable to examine associations between average caregiver behavior and average child behavior. These analyses, however, should not be viewed as evidence of causal relationships, but only as indications of associations among ongoing classroom processes.

Process relationships were examined by computing correlations between children's behaviors (combination teacher-directed and free-play) and both teachers' and aides' behaviors. Additional correlations between child and adult behavior were also obtained by caregivers' level of education. The results are summarized in Tables 4.1 to 4.3. Actual correlations are presented in Tables A.16 to A.21 in Appendix A.

Overall, the correlational pattern obtained suggests that the AFI and CFI observations represent a continuing classroom process. The correlations do not create a clearly defined picture of the relationships between child and adult behavior, but rather point toward certain configurations of classroom process that are evident in both behavior codes. For example, caregivers who interacted more with the children (more instructing, social interaction and management behavior) were found in classes where children were more involved (giving orders, receiving orders, receiving information). On the other

Figure 4.1

SUMMARY OF ADULT FOCUS AND CHILD FOCUS BEHAVIORS TEACHER-DIRECTED AND FREE-PLAY ACTIVITIES COMBINED

CAREGIVERS	CHILDREN
For Both Teachers and Aides:	
• more social interaction	• more participating in closed group activity • more receiving information, help
For Teachers Only	
• more social interaction	• more giving opinions • less wandering
• more management behavior	• less giving orders
• more center activity	• more giving orders
For Aides Only:	
• more social interaction	• less monitoring environment • less participating in open group activity
• more management behvaior	• more giving opinions • more receiving orders
• more adult activity	• more doing nothing, wandering • less receiving information • less receiving orders
• more observing	• more participating in open group activity • less receiving information, help

Figure 4.2

SUMMARY OF ADULT FOCUS AND CHILD FOCUS BEHAVIORS
TEACHER-DIRECTED ACTIVITIY

CAREGIVERS	CHILDREN
For Both Teachers and Aides:	
• more social interaction	• less monitoring environment • less doing nothing, wandering
• more adult activity	• less receiving information, help
For Aides Only:	
• more social interaction	• more participating in closed group activity
• more management behavior	• more participating in open group activity
• more adult activity	• more doing nothing, wandering
• more center activity	• more receiving information, help
• more instructing	• more participating in closed group activity • less doing nothing, wandering
For Teachers Only	
• more social interaction	• more receiving information, help
• more management behavior	• more participating in open group activity
• more observing	• less participating in open group activity
• more adult activity	• more monitoring environment • less participating in open group activity • less participating in closed group activity • more giving opinions

Figure 4.3

SUMMARY OF ADULT FOCUS AND CHILD FOCUS BEHAVIORS
FREE-PLAY

CAREGIVERS	CHILDREN
For Both Teachers and Aides:	
• more instructing	• more receiving information, help
For Teachers Only	
• more social interaction	• less doing nothing, wandering • less participating in open group activity • more giving opinions • more receiving information, help
• more management behavior	• more giving opinions • less giving orders
• more center activity	• less giving opinions • more giving orders • less receiving general comment
For Aides Only:	
• more social interaction	• less monitoring environment • less participating in open group activity • more participating in closed group activity • more receiving orders
• more management behaior	• less participating in open group activity • more giving opinions • more receiving orders. • more receiving information, help
• more adult activity	• more receiving general comment • less receiving information, help
• more observing	• less receiving information, help
• more instructing	• more receiving orders
• more center activity	• less participating in closed group activity

hand, caregivers who interacted with children less often (more adult activity, center activity and observing) were associated with children who were less involved (more monitoring and wandering, less giving and receiving). The correlations also showed a consistent relationship between children's participation in closed group activities and involved caregivers and, conversely, between children's participation in open group activities and uninvolved caregivers.

Relationships Between CDA and Adult-Focus Variables

To examine CDA/AFI relationships, correlations were computed between the major constructs defined for each instrument. The results (Table 4.4) show a consistent pattern for both teachers and aides. Caregivers who received higher CDA ratings engaged in more social interaction and center activity and less observing and adult activity. CDA ratings, however, were not significantly related to either instructing or management behavior in the APS sample.

Since the CDA Checklist was recorded immediately after the Adult-Focus Observations were made, the correlational pattern is not surprising. Many of the items on the CDA Checklist refer to specific activities concerning interaction with children or classroom management. Caregivers who did these things should have been recorded more often as praising, responding, comforting, or engaging in center-related activity. Thus to a large extent the CDA variables represent another measure of caregiver behavior, with a greater emphasis on skills than the Adult-Focus Instrument contains.

Relationships Between CDA and Child-Focus Variables

Correlations between the CDA variables and the CFI variables are shown in Table 4.5. Several consistent

Table 4.4

PEARSON CORRELATIONS FOR CDA AND ADULT FOCUS VARIABLES

	CDA Rating	Child Orientation	Physical Activity	Cog./Lang. Development	Self-Concept	Self Help Skills	Social Behavior	Classroom Management	Safety Awareness	Manages Activities Well	Resources	Gross Motor Toys	Creative Materials	Arranges Room Well	Physical Environment	Maintains Safe Classroom	Maintains Sanitary Classroom
Instructs																	
Teachers						.23			.20								
Aides									-.34								
Adult Activity																	
Teacher	-.42**	-.32*	-.22	-.33*	-.32*	-.26		-.23	-.20	-.22	-.44**	-.21	-.50*	.40*	-.44**	-.34*	-.33*
Aides	-.54**	-.50**	-.29	-.62**	-.42*			-.48**	-.25	.51**	-.50**	-.48**	-.45*	-.41*	-.51**		-.44*
Center Activity																	
Teacher	.39*	.33*		.38*	.29		.23	.35*	.38*	.29	.33*		.31*	.38*	.28		.27
Aides	.24					.63*					.36*	.47**	.35*			.27	
Observes																	
Teachers	-.44**	-.46**	-.21	-.54**	-.41**		-.31*	-.60**	-.60**	-.52**	-.20		-.21	-.27	-.20		-.24
Aides	.30		-.31		-.24	-.35					-.40*	-.31	-.45*	-.29		-.23	
Mgmnt. Behavior																	
Teachers		.27															
Aides		.34	.26					.33		.38*							
Social Interact																	
Teachers	.22	.36*	.20	.40*	.39*		.23	.38*	.37*	.34*							
Aides	.31	.30		.41*	.39*					.23	.32		.40*	.33	.42*	.33	.42*

*p<.05
**p<.01
Correlations are reported for p<.15 only.

Table 4.5

PEARSON CORRELATIONS FOR CDA AND CHILD FOCUS VARIABLES

(N=30)

	CDA Rating	Child Orientation	Physical Activity	Cog./Lang. Development	Self-Concept	Self Help Skills	Social Behavior	Classroom Management	Safety Awareness	Manages Activities Well	Resources	Gross Motor Toys	Creative Materials	Arranges Room Well	Physical Environment	Maintains Safe Classroom	Maintains Sanitary Classroom
Mon. Envir.	-.51**	-.[illegible]8**	-.22	-.51**	-.66**	-.21	-.50*	-.36*	-.33*	-.33*	-.36*	-	-.32*	-.48**	-.27	-	-.56**
Free Play	-.54**	-.56**		-.45**	-.58**	-.30*	-.57	-.36*	-.30*	-.34	-.47	-	-.45**	-.56**	-.23	-	-.52**
Teacher Directed	-.47**	-.61**	-.23	-.55**	-.71**	-.22	-.47**	-.35*	-.32*	-.32*	-.26	-	-.19	-.39*	-.35*	-	-.61*
Wanders	-.37*	-.32*	-.23	-.37*	-.31*	-	-.26	-.42**	-	-.51**	-.30*	42**	-	-.30*	-	.21	-.27
Free Play	-	-	-	-	-	.26	-	-.35*	-	-.42**	-	-.19	-	-	-	-	-
Teacher Directed	-.40**	-.39*	-	-.46**	-.33*	-.19	-.20	-.26	-	-.32*	-.39*	-.48**	-.30*	-.26	-	-	-.27
Gives Opinions	-	-	.20	-	-	-	-	.22	-	.23	-.20	-	-.24	-	-	-	-
Free Play	-	-	.21	-	-	-	-	-	-	-	-.19	-	-.24	-	-	-	-
Teacher Directed	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-.21	-	-.26
Gives Orders	-	-	-	-	-	-	-	.22	.24	.19	-	-	-	-	-	-	-
Free Play	-	-	-	-	-	-.24	-	-	-	-	.21	.28	-	-	-	-	-
Teacher Directed	-	-.25	-	-.29	-.25	-	-	-	-	-	-	-	-	-	-	-	-
Receives Orders	-	-	-	-	-	-	-	-	-.20	-	-	.36*	-	-	-	-.33*	.32*
Free Play	-	-	-	-	-	-	-	-	-	-	-	.35*	-	-	-	-.31*	-
Teacher Directed	-	-	-	.19	-	-	-	-	-.21	-	-	-	-	-	-	-.20	.37
Receives Info.	.41**	.32*	.22	.36*	.29	-	-	.23	-	.29	.51**	.57**	.34*	.45**	-	-.31*	.21
Free Play	-	-	.22	-	-	-	-	-	-	.19	-	.24	-	-	-.22	-	-
Teacher Directed	.48**	.44**	-	.45**	.39*	.35*	.25	.26	-	.30	.50**	.48**	.36*	.48**	-	-	.34*
Receives General	-.33**	-.20	-	-	-	-.46**	-.19	-	-	-	-.54**	-.46**	-.52**	-.42**	-	-	-
Free Play	-.42**	-.35*	-	-.24	-.29	-.57**	-.32*	-.22	-.30*	-	-.45**	-.27	-.50**	-.38*	-.24	-	-.28
Teacher Directed	.26	-	-	-	-	-.26	-	-	-	-	-.39**	-.31*	-.29	-.43**	-	-	-.20
Open	-	-	-	-	-	.19	-	-	-	-	-	-	-	-	.27	.23	.19
Free Play	-	-	-	-	-	.25	-	-	-	-	-	-	-	.22	.21	-	-
Teacher Directed	-.33*	-	-.21	-	-	-	-	-	.19	-	-.60**	-.59**	-.41**	-.59**	-	.23	-.32*
Structure	.33*	.31*	-	.34*	.31*	-	.21	-	-	.20	.37*	-	.42**	.37*	-	-	.26
Free Play	-	-	-	.25	-	-	-	-	-	-	-	-.19	.19	-	-	-	-
Teacher Directed	.47**	.37*	.25	.34*	.37*	.24	.23	-	-	-	.61**	.44**	.58**	.56**	-	-	-

*p<.05
**p<.01
Correlations are reported for p<.15 only.

patterns are present. Caregivers with higher CDA ratings were found in classrooms in which children monitored the environment or wandered frequently and received more information and help. These relationships were especially strong during teacher-directed activities, when the caregiver might be expected to be in more direct contact with children. The CDA ratings also were associated with children's receiving fewer general comments and with children's participating in structured rather than open activities.*

Even though the CDA and Child-Focus Instruments were not recorded at the same time, the CDA/CFI relationships suggest that, as with the AFI, the behaviors recorded represent an on-going classroom process. Children in classes with higher CDA ratings tended to receive more from the caregiver and to be more involved, particularly in a structured activity. The effect of caregivers' skills and behaviors is less evident during free-play situations, except on children's monitoring of the environment and receiving general comments. The fact that these relationships are evident despite the difference in observation times helps to confirm the validity of both instruments.

---

*It should be emphasized that no attempt has been made to classify either adult or child behaviors as "good" or "bad." Individual variables may be good or bad depending on the overall situation and the perspective of the person making the judgment.

CHAPTER FIVE: ANALYSES OF CLASSROOM PROCESS AND CHILDREN'S COGNITIVE GAINS

Relationships between Adult-Focus Variables and Gain Scores

The relationships between the adult-focus variables and children's generalized cognitive gain scores were examined for both teachers and aides. Correlations (Table 5.1) indicate that in the APS study aide behaviors were more strongly associated with children's cognitive gains than were teacher behaviors. The AFI/policy variable analyses also suggested that aides were more influenced than teachers by the policy variables, especially by classroom composition (see Chapter Three). The effects of relevant policy variables were therefore removed through partial correlations. The resulting AFI/PSI correlations for aides were much less significant; the AFI/PPVT correlations for both aides and teachers, however, remained stable.

This finding is of particular interest if one is willing to place a causal "chain-of effects" model on these data. Suppose that interest lies with understanding the relationships between adult behavior and child outcomes. It is known, however, that caregiver qualifications and classroom composition are associated with both of these measures. If one assumes that these policy variables are in a sense "external inputs" to the system that may mediate between process and outcome, then by first partialing out these effects, it is possible to determine if classroom composition and caregiver qualifications are the determinants of child outcomes or if it is really process that influences cognitive gain scores.

In the context of this model, then, these findings suggest that the effects of adult (aide) behavior on PSI gain scores may be attributable to the prevailing classroom composition (group size, staff/child ratio), but that PPVT

Table 5.1

PEARSON CORRELATIONS FOR ADULT FOCUS VARIABLES AND COGITIVIE GAIN SCORES

	PSI Gains	PSI Gains Removing Group Score & S/C Ratio	PPVT Gains	PPVT Gains Removing Previous Day Care Experience	PPVT Gains Removing Group Size, S/C Ratio, Previous Day Care Experience
Instructs					
Teachers (N-30)					
Aides (N=22)			.28	.27	.31
Adult Activity					
Teachers					-.32
Aides	-.51**	-.24	-.32	-.28	-.50*
Center Activity					
Teachers		.27		.20	.20
Aides				-.40*	-.38*
Observes					
Teachers	-.27	-.21			
Aides	-.25				
Mgmt. Behavior					
Teachers	.20		-.20	-.20	-.21
Aides			-.35*	.42*	.49*
Social Interact.					
Teachers		-.25			
Aides	.44**	.25	-.45*	.44*	.56**
Negative					
Teachers		.25	-.30*		-.20
Aides			-.26*		-.24
Positive					
Teachers			.22	.21	.21
Aides	.25		.38*	.33	.40*

*$p<.05$
**$p<.01$
Correlations are reported for $p<.15$ only.

gain scores are more likely to be associated with adult behavior. PPVT gain scores were not initially as influenced by classroom composition, and the pattern of correlations remained stable when the classroom composition variables and PREVIOUS DAY CARE EXPERIENCE were first partialed out. In fact the correlations, especially for aides, tended to increase after the effects of the composition variables were removed. If one accepts this chain-of-effects model, these results suggest that classroom composition rather than caregiver behavior is related to PSI gain scores, and that caregiver behavior rather than qualifications or classroom composition is associated with PPVT gain scores.

The correlational pattern indicates that aides who showed more social interaction and management behavior were associated with higher PPVT gain scores, while aides who were more often involved in adult or center-related activities were associated with lower PPVT gain scores. These relationships, however, did not hold for lead teachers: teachers who evidenced less management behavior and more center activity were associated with greater PPVT gains. The inconsistent findings for teachers and aides are not easily explained and are not replicated in the 57-center study. This provides evidence that the reassignments made for the experimental design may have caused the anomalous results (see Chapter Three).

Regression analyses (Tables A.22 and A.23, Appendix A) also suggest that aide behaviors were not strong predictors of PSI gains after the effects of GROUP SIZE and STAFF/CHILD RATIO were removed. The regressions did indicate, however, that for lead teachers SOCIAL INTERACTION and OBSERVES were significant predictors of PSI GAIN. In fact, the teacher-focus social interaction variable became a stronger predictor after the classroom composition variables were considered. GROUP SIZE, however, remained the best predictor of PSI GAIN for both teachers and aides. Regression results for the

PPVT showed that MANAGEMENT BEHAVIOR for lead teachers (especially CORRECTS) and SOCIAL INTERACTION for aides were the best predictors of PPVT GAIN.

The correlation and regression results together suggest that caregiver behavior, particularly SOCIAL INTERACTION, is associated with children's cognitive gains in the APS sample. Both PREVIOUS DAY CARE EXPERIENCE and the caregiver behavior variables were good predictors of PPVT GAIN, while GROUP SIZE remained the best predictor of PSI GAIN. Although aide behaviors appeared to be related to PSI gains, the effects disappeared after group size was considered; however, teacher behaviors, which were less influenced by group size, continued to be associated with PSI gain scores.

Relationships between Child-Focus Variables and Gain Scores

Correlations between the child-focus variables and children's cognitive gain scores are presented in Table 5.2. The pattern provides evidence that children's behaviors were more closely associated with changes in PPVT scores than with changes in PSI scores. After the effects of classroom composition were removed, only two CFI variables--GIVES OPINIONS during teacher-directed activities and GIVES ORDERS during free play--were significantly correlated with PSI GAIN. Although removing PREVIOUS DAY CARE EXPERIENCE from the CFI/PPVT correlations generally decreased the values obtained, a number of significant relationships remained. PARTICIPATES IN STRUCTURED ACTIVITY (for example, games with rules, working puzzles or coloring) was positively and significantly correlated with PPVT GAIN while MONITORS ENVIRONMENT, WANDERS and PARTICIPATES IN OPEN ACTIVITY (for example, riding bikes, drawing pictures or building with blocks) were negatively correlated with PPVT GAIN. GIVES ORDERS and RECEIVES INFORMATION, especially during free play, were also positively associated with PPVT GAIN. Including GROUP SIZE and STAFF/CHILD RATIO in the PPVT

Table 5.2

PEARSON CORRELATIONS FOR CHILD FOCUS VARIABLES AND COGNITIVE GAIN SCORES

(N=30)

	PSI Gains	PSI Gains Adjusting for Group Size and S/C Ratio	PPVT Gains	PPVT Gains Adjusting for Previous Day Care Experience	PPVT Gains Adjusting for Prev. Day Care Experience, group Size & S/C Ratio
Mon. Envir.					
Combined	-.23		-.32		-.21
Teacher Directed		.21	-.19		
Free Play	-.21		-.42**	-.33*	-.37*
Wanders					
Combined			-.49**	-.32*	-.32*
Teacher Directed	-.28		-.36*	-.21	-.22
Free Play			-.33*	-.24	-.24
Gives Opinions					
Combined	.19				
Teacher Directed	.57**	.51**	.20	.26	.27
Free Play		-.26			
Gives Orders					
Combined		.25	.28	.31*	.34*
Teacher Directed					
Free Play		.41*	.40**	.44**	.45**
Receives Orders					
Combined					
Teacher Directed	-.22	-.30			
Free Play	.33*				
Receives Info.					
Combined	.46**		.32*		.19
Teacher Directed	.26		.20		
Free Play	.21		.32*	.32*	.33*
Receives General					
Combined					
Teacher Directed	.21	.22			
Free Play					
Open Activity.					
Combined	-.36*		-.41**	-.41**	-.45**
Teacher Directed			-.34*	-.21	-.22
Free Play				-.20	-.19
Structured Act.					
Combined	.23		.50**	.47**	.50**
Teacher Directed	.25		.44**	.40*	.43**
Free Play			.33*	.36*	.36*

*p<.05
**p<.01
Correlations are reported for p<.15 only.

partial correlations did not change the correlational patterns.

In the initial regression analyses, child-focus variables were used to predict gain scores by type of observation (teacher-directed activity, free play, combined). Regressions also were computed first taking into account the effects of selected policy variables. The results (Tables A.24 to A.26, Appendix A) corroborated the correlational findings and emphasized that PPVT gain scores were especially influenced by the free-play variables.

After the effects of classroom structure were taken into consideration, the only child-focus variable that remained a significant predictor of PSI GAIN was GIVES OPINIONS during teacher-directed activities (p=.058). PPVT gains, however, were influenced by several free-play variables, even after the effects of PREVIOUS DAY CARE EXPERIENCE were removed from the model. GIVES ORDERS, PARTICIPATES IN OPEN ACTIVITY, MONITORS ENVIRONMENT and RECEIVES INFORMATION each significantly predicted PPVT GAIN.

While it is difficult to interpret the regression models that include the isolated significant predictors, two interesting patterns do emerge. First, children's behaviors were more strongly related to PPVT gain scores than to PSI gain scores in the Atlanta Public Schools day care centers. Second, child behavior during teacher-directed activities (especially GIVES OPINIONS) was more closely associated with PSI gain scores, whereas child behavior in free-play situations seemed to be more strongly related to PPVT changes.

Relationships between CDA Variables and Gain Scores

Simple and partial correlations for the CDA variables and PSI and PPVT gain scores are shown in Table 5.3.

Table 5.3

PEARSON CORRELATIONS FOR CDA VARIABLES AND COGNITIVE GAIN SCORES

	PSI	PSI Controlling for Group Size and S/C Ratio	PPVT	PPVT controlling for Previous Day Care Experience	PPVT Controlling for Previous Day Care Experience Group Size and Staff/Child Ratio
CDA Rating	.33*		.55	.42	.45
Child Orientation	.19		.44	.30	.31
Encourages Active Play			.31		
Encourages Cognite/ Language Development	.26		.40	.30	.31
Encourages good Self-Concept			.47	.33	.35
Encourages Self Help					
Encourages Soc. Behav.			.46	.32	.32
Classroom Management	.25		.43**	.33*	.33*
Encourages Safety	.27				
Manages Class Act. Well	.21		.52**	.41**	.42**
Resources	.43**	.29**	.56**	.45**	.48**
Provides Gross Motor Materials.	.37*	.28*	.49**	.36**	.37*
Provides Createive Play Materials	.49**	.37*	.45**	.41**	.44**
Arranges Classroom Well	.24		.52**	.38*	.41*
Physical Environment	-.04		.08		
Maintains Safe Classrm.			-.12	-.07	
Maintains Sanitary Classroom			.24		

*p<.05
**p<.01
Correlations are reported for p<.15 only.

Analyses of the CDA variables and the policy variables (see Chapter Three) indicated that CDA ratings were influenced by both staff/child ratio and caregivers' previous day care experience. Although several CDA variables had significant or nearly significant correlations with PSI gain scores, the effects were greatly diminished if GROUP SIZE and STAFF/CHILD RATIO were first considered; only one correlation, CREATIVE PLAY MATERIALS with PSI GAIN, then remained significant. The overall CDA rating was significantly correlated with PSI GAIN, but the correlation became insignificant if the effects of the classroom composition variables were first removed. (This diminished effect is largely due to the correlations between CDA measures and classroom composition measures.) Thus it apears difficult to separate out the effects of the CDA ratings and classroom composition measures upon PSI gains.

The relationship between CDA ratings and PPVT gain scores was a much stronger one. All but five of the CDA constructs were significantly correlated with PPVT GAIN before policy variables were considered. Controlling for PREVIOUS DAY CARE EXPERIENCE (or PREVIOUS DAY CARE EXPERIENCE, GROUP SIZE and STAFF/CHILD RATIO) diminished relationships, especially for CHILD ORIENTATION. Nonetheless, a number of significant partial correlations remained.

Regression results (Tables A.27 and A.28, Appendix A) indicated that the CDA variables were not strong predictors of PSI gain scores but that the overall CDA rating was a better predictor of PPVT gains than were any of the policy or process variables. Its significance, however, was greatly reduced if PREVIOUS DAY CARE EXPERIENCE was first entered into the regression model. Examining the regression coefficients, it is clear that the two variables are multi-collinear and that their effects can not be separated. The evidence suggests, however, that caregivers who work well with children, who have had experience working with children in other day care settings and who direct their attention to the needs of the children are associated with higher PPVT gain scores, but not necessarily with higher PSI gain scores.

CHAPTER SIX: CONCLUSIONS

The analyses reported above, which are based on data for approximately 260 children in eight Atlanta Public Schools day care centers, have highlighted several relationships between classroom composition and caregiver qualifications, adult and child behaviors, and children's acquisition of skills and knowledge as measured by the PSI and PPVT. Most important, these results reaffirm the theory that classrooms in which small numbers of children and caregivers interact with each other are associated with better care for children. Moreover, several aspects of caregiver qualifications, most notably previous day care experience, are not only related to the degree of caregiver interactiveness but are also associated with higher cognitive gain scores.

In addition to these broad results, the replication analysis in the Atlanta Public Schools suggests that a distinction should be made between teachers and aides. In the APS study, aides and lead teachers behaved differently and were differentially affected by variations in classroom composition. Although on the surface this result may not be surprising, it helps to provide insight into classroom dynamics and perceived job roles. For example, the APS aides were much more interactive in smaller classes with higher ratios; when the classes were too large and the ratio was lower, the aides tended to observe classroom activities. On the other hand, lead teacher behavior was not significantly influenced by these variations; the APS lead teachers appeared to cope with larger groups in much the same manner as they did with smaller ones.

The amount of previous day care experience a caregiver had also appeared to be a key determinant of better care for children. Caregivers with more experience tended to be more interactive and to receive higher CDA ratings; moreover, they were able to keep children actively involved in classroom activities. This active integration into the classroom is reflected in higher PPVT gain scores. Thus, it appears that more experienced caregivers are better equipped to manage a classroom well and to facilitate those environments that are associated with better child care.

It is important to note that, except for the findings for aides, the results reported here parallel those found in the 57-center study. Both studies indicate that a small-scale classroom environment and skilled, experienced caregivers benefit children in day care. However, in the APS study, which had little variation in specialization, caregiver quality is reflected in the variable PREVIOUS DAY CARE EXPERIENCE, while in the overall study, in which there was little variation in previous day care experience, the variable SPECIALIZATION appears to be associated with caregiver quality. Thus, there is some evidence that caregivers who have a continuing interest in day care and children, which is reflected in their either seeking new jobs in the field or obtaining intensive, specialized preparation related to day care, provide a btter environment in which young children can thrive.

References

Bache, W.L. "Comparing Alternative Measures of Classroom Composition." *National Day Care Study Measurements and Methods*. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980a.

Connell, D.C. "The Child-Focus Observation Effects Analysis." *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Goodrich, N.N. "The Atlanta Public Schools Day Care Experiment." *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980a.

Goodrich, N.N. "The Effects of Day Care in Eight Atlanta Public Schools Day Care Centers." *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980b.

Goodrich, N.N. "An Analysis of the CDA Checklist Data." *National Day Care Study Measurements and Methods*. Final Report of the National Day Care Study, Volume IV-B. Cambridge, Mass.: Abt Associates Inc., 1980c.

Goodrich, R.L., and Singer, J.D. "Cognitive Change in the NDCS." *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Goodson, B.D. "The Adult-Focus Observation Effects Analysis." *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

Singer, J.D. "Classroom Process-Child Outcome (Process-Outcome) Analysis." *National Day Care Study Effects Analyses*. Final Report of the National Day Care Study, Volume IV-C. Cambridge, Mass.: Abt Associates Inc., 1980.

APPENDIX A

Table A.1

POLICY VARIABLE/PSI GAIN SCORE REGRESSION RESULTS CHILD LEVEL VS. CLASS LEVEL

Independent Variable	Child Level (N=265)						Class Level (N=30)					
	r	B	SE_B	F	P_F	R^2	r	B	SE_B	F	P_F	R^2
Years of Education	.07	.323	.267	1.47	.22	.01	.14	.339	.446	.53	.52	.02
Staff/Child Ratio	.21	28.324	8.338	11.54	.001	.04	.37	28.759	13.915	4.27	.05	.13
Group Size	-.35	-.310	.051	37.01	.000	.12	-.62	-.310	.074	17.30	.01	.38
Number of Caregivers	-.10	-.561	.350	2.56	.11	.01	-.17	-.565	.609	.86	.64	.03
Center Experience	.02	.068	.197	.12	.73	.00	.03	.060	.344	.03	.86	.00
Previous Day Care Experience	-.002	(not calculatable)					.009	(not calculatable)				
Specialization	.08	2.067	1.540	1.80	.18	.01	.15	2.103	2.659	.62	.56	.02
Level of Education	.05	.316	.354	.80	.62	.00	.08	.255	.572	.20	.66	.01
Group Size Staff/Child Ratio	-.35 .21	-.288 20.973	.051 7.996	31.82 6.88	.000 .009	.15	-.62 .37	-.287 20.982	.073 11.449	15.62 3.36	-.00 .075	.45
Group Size Center Experience	-.35 .02	-.343 .421	.053 .191	42.30 4.89	.000 .03	.14	 .03	-.342 .415	.976 .276	20.28 2.25	.00 -.13	.43
Group Size Previous Day Care Experience	-.35 -.002	-.349 -.692	.054 .314	42.41 4.86	.000 .03	.14	-.62 -.009	-.349 -.706	.077 .456	20.57 2.40	.00 .13	.43
Group Size Specialization	-.35 .08	-.314 2.537	.051 1.443	88.33 3.09	.000 .08	.13	-.62 .14	-.314 2.545	.974 2.099	18.06 1.47	.00 .23	.41
Group Size Years of Education	-.35 .07	-.306 .173	.051 .252	35.74 .47	.000 .50	.13	-.62 .14	-.306 .180	.076 .377	16.24 0.23	.00 .64	.39
Group Size Center Experience Previous Day Care Experience	-.35 .02 -.002	-.395 .499 -.819	.056 .191 .314	50.03 6.82 6.79	.000 .009 .01	.16	-.62 .03 .009	-.395 .496 .837	.078 .267 .442	25.80 3.44 3.58	.00 .07 .07	.50
Group Size Staff/Child Ratio Specialization	-.35 .21 .08	-.290 25.486 3.571	.051 8.132 1.457	32.86 9.82 6.01	.000 .002 .01	.17	-.62 .37 .15	-.288 25.536 3.571	.070 11.310 2.008	17.02 5.10 3.16	.001 .03 .08	.51
Group Size Staff/Child Ratio Level of Education	-.35 .21 .05	-.283 21.723 .215	.052 8.090 .334	30.17 7.21 .41	.000 .008 .53	.15	-.62 .37 .08	-.282 11.814 .449	.074 3.664 .459	14.60	.00 .07 .51	.46

Table A.2

POLICY VARIABLE/PPVT GAIN SCORE REGRESSIONS RESULTS
CHILD LEVEL VS. CLASS LEVEL

Independent Variable	Child Level (N=265)						Class Level (N=30)					
	r	B	SE_B	F	P_F	R^2	r	B	SE_B	F	P_F	R^2
Years of Education	.04	.410	.565	.53	.52	.01	.14	.504	.711	.501	.47	.02
Staff/Child Ratio	.06	-18.354	-18.344	1.00	.32	.04	.37	-17.679	23.014	4.596	.04	.02
Group Size	-.09	- .164	.117	1.97	.16	.12	-.62	-.180	.144	1.538	.22	.05
Number of Caregivers	-.05	- .577	.748	.59	.55	.01	-.17	-.631	.942	.45	.51	.02
Center Experience	.04	.284	.406	.49	.51	.00	.03	-.315	.509	.38	.54	.01
Previous Day Care Experience	.21	2.315	.663	12.19	.001	.04	.51	2.306	.740	9.72	.00	.26
Specialization	.11	5.548	3.181	3.04	.08	.01	.26	5.499	3.939	1.95	.17	.07
Level of Education	.08	.916	.739	1.54	.21	.01	.25	1.160	.838	1.92	.17	.06
Previous Day Care Experience	.21	2.300	.664	11.99	.001	.05	.51	2.291	9.746	9.42	.01	.27
Years of Education	.04	.342	.554	.02	.54		.25	.450	.624	.05	.82	
Previous Day Care Experience	.21	2.292	.687	11.12	.001	.04	(not calculatable)					
Staff/Child Ratio	.06	-2.358	18.622	.02	.89							
Previous Day Care Experience	.21	2.243	.702	10.21	.002	.05	.51	2.199	.794	7.67	.01	.26
Group Size	-.09	.038	.121	.10	.75		-.23	- .036	.137	.17	.69	
Previous Day Care Experience	.21	2.349	.664	12.51	.001	.05	.51	2.344	.744	9.94	.00	.28
Center Experience	-.04	-.364	.398	.84	.64		-.12	-.396	.444	.80	.62	
Previous Day Care Experience	.21	2.183	.671	10.60	.002	.05	.51	2.177	.746	8.51	.01	.29
Specialization	-.11	3.922	3.163	1.54	.21		.26	3.868	3.542	1.19	.28	
Previous Day Care Experience	.21	2.198	.670	10.76	.002	.06	.51	2.194	.740	8.79	.01	.33
Center Experience	-.04	-.518	.410	2.30	.13		-.12	-.550	.451	1.49	.23	
Specialization	.01	4.947	3.261	1.60	.20		.26	4.965	3.624	1.89	.18	
Previous Day Care Experience	.21	2.335	.666	12.30	.001	.05	.51	2.328	.752	9.60	.00	.29
Center Experience	-.04	-.351	.399	.77	.62		-.12	-.379	.449	.71	.59	
Years of Education	.04	.314	.555	.32	.58		.13	.419	.628	.44	.52	
Previous Day Care Expereince	.21	2.279	.664	11.80	.001	.05	.51	2.243	.729	9.47	.01	.31
Level of Education	.08	.793	.725	1.20	.27		.25	1.020	.736	1.92	.17	
Previous Day Care Experience	.21	2.317	.664	12.19	.001	.05	.51	2.284	.726	9.90	.004	.34
Level of Education	.08	1.032	.749	1.90	.17		.25	1.149	.740	2.41	.13	
Center Experience	-.04	- .510	.411	1.54	.21		-.12	-.500	.438	1.30	.26	

Table A.3

POLICY VARIABLE/PSI GAIN SCORE REGRESSION RESULTS - CHILD LEVEL LOGGED VS. UNLOGGED CLASS STRUCTURE VARIABLES

(N=265)

Independent Variable	Unlogged						Logged					
	r	B	SE_B	F	P_F	R^2	r	B	SE_B	F	P_F	R^2
Group Size	-.39	-.350	.055	40.11	.000	.15	-.31	-11.705	2.429	23.23	.000	.09
Group Size	-.39	-.335	.005	37.49	.000	.17	-.31	-10.985	2.384	21.23	.000	.14
Staff/Child Ratio	.19	20.868	8.786	5.64	.02		.24	10.230	3.023	11.45	.001	
Group Size	-.39	-.388	.058	45.46	.000	.17	-.31	-14.744	2.715	29.48	.000	.12
Previous Day Care Experience	.01	-.698	.343	4.14	.04		.01	- .910	.379	5.78	.02	
Group Size	-.39	-.376	.057	43.95	.000	.16	-.31	-12.464	2.511	24.65	.000	.10
Center Experience	-.01	.343	.203	2.85	.089		-.01	.248	.210	1.39	.24	
Group Size	-.39	-.349	.055	40.08	.000	.15		(not calculatable)				
Specialization	.05	.655	1.670	.15	.70							
Group Size	-.39	-.363	.059	38.68	.000	.18	-.31	-13.093	2.737	22.87	.000	.15
Staff/Child Ratio	.19	16.438	9.479	3.01	.08		.24	8.862	3.140	7.96	.005	
Previous Day Care Experience	.01	-.456	.369	1.53	.21		.01	- .603	.388	2.41	.12	
Group Size	-.39	-.428	.060	50.80	.000	.19	-.31	-16.550	2.872	33.22	.000	.13
Previous Day Care Experience	.01	-.845	.347	5.92	.01		-.01	- 1.088	.389	7.84	.006	
Center Experience	-.01	.441	.205	4.62	.03			.396	.214	3.43	.06	
Group Size	-.39	-.363	.056	41.48	.000	.19	-.31	-11.942	2.453	23.70	.000	.15
Staff/Child Ratio	.19	22.160	8.762	6.40	.01		.24	10.724	3.030	12.53	.00	
Center Experience	-.01	.383	.201	3.61	.05		-.01	.323	.206	2.46	.11	
Group Size	-.39	-.329	.055	35.77	.000	.18	-.31	-10.968	2.376	21.31	.000	.15
Staff/Child Ratio	.19	23.437	9.109	6.62	.010		.24	11.182	3.072	13.25	.000	
Specialization	.05	1.821	1.710	1.13	.29		.05	2.716	1.709	2.52	.11	
Group Size	-.39	-.387	.058	45.26	.000	.17	-.31	-15.192	2.727	31.04	.000	.12
Previous Day Care Experience	.01	-.744	.349	4.55	.032		.01	-1.031	.387	7.10	.008	
Specialization	.51	1.274	1.682	0.57	.54		.05	2.506	1.740	2.07	.15	
Group Size	-.39	(not calculatable)					-.31	-12.382	2.518	24.18	.000	.10
Center Experience	-.01						-.01	.215	.218	.97	.67	
Specialization	.05						.05	1.051	1.782	.35	.56	
Group Size	-.39	(not calculatable)					-.31	-10.940	2.409	20.63	.000	.14
Staff/Child Ratio	.19						.24	10.226	3.030	11.39	.001	
Years of Education	.05						.05	.042	.281	.02	.88	
Group Size	-.39	-.375	.057	43.11	.000	.16	-.31	-12.400	2.533	23.97	.000	.10
Center Experience	-.01	.344	.204	2.85	.09		-.01	.250	.211	1.40	.24	
Years of Education	.05	.041	.277	.02	.88		.05		.288	.05	.82	

Table A.4

POLICY VARIABLE/PPVT GAIN SCORE REGRESSION RESULTS - CHILD LEVEL LOGGED VS. UNLOGGED CLASS STRUCTURE VARIABLES

(N=228)

Independent Variable	Unlogged						Logged					
	r	B	SE_B	F	P_F	R^2	r	B	SE_B	F	P_F	R^2
Group Size	-.11	-.069	.128	.29	.60	.05	-.14	-3.492	5.867	.35	.56	.05
Previous Day Care Experience	.23	2.386	.765	9.74	.002		.23	2.292	.818	7.85	.006	
Group Size	-.11	.081	.132	.31	.58	.05	-.14	-3.878	6.018	.41	.53	.05
Staff/Child Ratio	-.08	-7.642	21.254	.13	.72		-.06	-2.072	6.904	.09	.76	
Previous Day Care Experience	.23	2.274	.828	7.55	.006		.23	2.220	.954	6.76	.01	
Group Size	-.11	-.042	.135	.10	.76	.05	-.14	-2.207	6.247	.12	.72	.05
Previous Day Care Experience	.23	2.483	.781	10.11	.002		.23	2.419	.845	8.19	.005	
Center Experience	-.03	-.292	.461	.04	.83		-.03	-.282	.465	.37	.55	
Previous Day Care Experience	.23	2.504	.766	10.68	.002	.05	.23	2.534	.743	11.64	.001	.05
Staff/Child Ratio	-.08	-5.789	20.659	.08	.76		-.06	-1.695	6.780	.06	.57	
Center Experience	-.03	-.347	.439	.63	.56		-.29	-.349	.440	.63	.57	
Group Size	-.11	-.068	.128	.28	.60	.06	-.14	-4.218	5.904	.51	.52	.06
Previous Day Care Experience	.23	2.253	.776	8.42	.004		.23	2.097	.838	6.27	.01	
Specialization	.11	3.729	3.745	.99	.68		.11	4.058	3.767	1.16	.28	

Table A.5

POLICY VARIABLE/ADULT FOCUS REGRESSION RESULTS - TEACHERS

INSTRUCTS
(N=29)

	r	B	SE_B	F	P_F	R^2
Total Day Care Experience	.26	.006	.004	1.98	.17	.07
Number of Caregivers	-.08	-.005	.011	.21	.65	
Education Level	.15	.007	.010	.60	.55	.10
Total Day Care Experience	.26	.006	.004	2.01	.17	
Staff/Child Ratio	-.20	.140	.154	.82	.62	.10
Total Day Care Experience	.26	.005	.004	1.70	.20	
Staff/Child Ratio	-.20	-.197	.161	1.49	.23	.14
Education Level	.15	.011	.010	1.32	.26	
Total Day Care Experience	.26	.005	.004	1.42	.24	

ADULT ACTIVITY
(N=29)

	r	B	SE_B	F	P_F	R^2
Number of Caregivers	-.36	-.021	.010	4.07	.05	.13
Staff/Child Ratio	-.32	-.273	.154	3.13	.08	.10
Education Level	-.15	-.008	.010	.64	.56	.02
Group Size	.001	-.002	.002	.75	.60	.13
Staff/Child Ratio	-.32	-.339	.173	3.85	.06	
Group Size	.001	-.002	.002	.82	.62	.14
Staff/Child Ratio	-.32	-.335	.175	3.64	.06	
Total Day Care Experience	.11	.002	.004	.27	.61	
Number of Caregivers	-.36	-.022	.011	4.20	.05	.16
Education Level	-.15	-.009	.009	.87	.64	
Number of Caregivers	-.36	-.022	.011	4.30	.05	.18
Total Day Care Experience	.11	.003	.004	.83	.63	
		-.009	.009	.53	.52	
Staff/Child Ratio	-.32	-.151	.182	.68	.58	.15
Number of Caregivers	-.36	-.015	.013	1.52	.23	
Staff/Child Ratio	-.32	-.266	.167	2.85	.10	.11
Total Day Care Experience	-.15	.002	.004	.18	.68	
Staff/Child Ratio	-.32	.132	.187	.50	.51	.17
Number of Caregivers	-.36	-.017	.013	1.67	.21	
Total Day Care Experience	.11	.003	.004	.37	.56	

Table A.5 (cont'd)

CENTER ACTIVITY
(N=29)

	r	B	SE_B	F	P_F	R^2
Education Level	.41	.058	.024	5.66	.02	.17
Number of Caregivers	.14	.025	.028	.79	.62	.19
Education Level	.41	.058	.025	5.55	.02	
Number of Caregivers	.14	.023	.028	.69	.58	.21
Total Day Care Experience	.14	.009	.011	.63	.56	
Education Level	.41	.058	.025	5.53	.02	
Staff/Child Ratio	.38	.647	.404	2.57	.12	.25
Education Level	.41	.046	.025	3.45	.07	
Staff/Child Ratio	.38	.695	.409	2.89	.10	.27
Education Level	.41	.044	.025	3.06	.09	
Total Day Care Experience	.14	.011	.010	1.10	.30	
Without Directors (N=21)						
Staff/Child Ratio	.48	1.25	.52	5.79	.02	.23
Staff/Child Ratio	.48	1.39	.47	8.58	.007	.41
Total Day Care Experience	.35	.021	.009	5.24	.03	
Staff/Child Ratio	.48	1.64	.49	7.68	.01	.41
Center Experience	.16	.020	.010	4.20	.05	
Previous Day Care Experience	.23	.024	.014	3.11	.09	

Table A.5 (cont'd)

OBSERVES
(N=29)

	r	B	SE_B	F	P_F	R^2
Education Level		-.054	.022	5.74	.02	.18
Group Size	.40	.006	.004	3.00	.09	.26
Education Level	-.40	-.043	.023	3.60	.07	
Number of Caregivers	.11	.018	.025	.51	.51	.19
Education Level	-.40	-.054	.023	5.74	.02	
Number of Caregivers	.11	.015	.026	.31	.59	.18
Total Day Care Experience	-.08	-.005	.010	.24	.63	
Education Level	-.40	-.052	.023	4.89	.03	
Group Size	.40	.008	.004	3.37	.07	.27
Education Level	-.40	-.046	.023	3.88	.06	
Staff/Child Ratio	-.14	.267	.405	.44	.52	
Staff/Child Ratio	-.14	-.262	.471	.30	.59	.20
Education Level	-.40	-.050	.024	4.16	.05	
Number of Caregivers	.11	.028	.031	.80	.62	

MANAGEMENT BEHAVIOR
(N=29)

	r	B	SE_B	F	P_F	R^2
Education Level	-.19	-.013	.013	1.06	.31	.04
Number of Caregivers	.15	.013	.015	.72	.59	.09
Total Day Care Experience	-.16	-.005	.006	.66	.57	
Education Level	-.19	-.013	.013	.98	.67	
Group Size	.15	.002	.002	.44	.52	.08
Education Level	-.19	-.010	.014	.52	.52	
Total Day Care Experience	-.16	-.005	.006	.74	.60	
Education Level	-.19	-.013	.013	.94	.66	.06
Total Day Care Experience	-.16	-.004	.006	.59	.55	

Table A.5 (cont'd)

SOCIAL INTERACTION

(N=29)

	r	B	SE_B	F	P_F	R^2
Group Size	-.21	-.003	.003	1.24	.27	.04
Number of Caregivers	-.30	-.027	.017	2.70	.11	.09
Staff/Child Ratio	-.21	-.271	.248	1.19	.28	.04
Group Size	-.21	-.005	.003	3.47	.07	.16
Staff/Child Ratio	-.21	-.490	.265	3.42	.07	
Number of Caregivers	-.30	-.028	.017	2.87	.10	.14
Education Level	.20	.018	.015	1.41	.24	
Number of Caregivers	-.30	-.027	.017	2.41	.13	.13
Total Day Care Experience	-.06	-.001	.007	.03	.85	
Education Level	.20	.014	.014	.96	.66	
Group Size	-.21	-.004	.003	1.68	.20	.21
Education Level	.20	.020	.015	1.60	.22	
Staff/Child Ratio	-.21	-.559	.268	1.60	.22	
Staff/Child Ratio	-.21	-.215	.307	.49	.50	.15
Education Level	.20	.021	.016	1.80	.19	
Number of Caregivers	-.30	-.020	.020	.99	.67	

Table A.6

POLICY VARIABLE/ADULT FOCUS REGRESSION RESULTS - AIDES

INSTRUCTS
(N=21)

	r	B	SE_B	F	P_F	R^2
Group Size	-.57	-.011	.004	9.15	.007	.32
Number of Caregivers	-.16	-.020	.028	.51	.51	.03
Staff/Child Ratio	.32	.588	.401	2.15	.16	.10
Number of Caregivers	-.16	-.091	.031	8.39	.009	.39
Staff/Child Ratio	.32	1.547	.475	10.61	.004	
Group Size	-.57	-.011	.004	9.02	.008	.33
Total Day Care Experience	.01	.003	.005	.28	.61	
Number of Caregivers	-.16	-.021	.029	.53	.52	.03
Total Day Care Experience		.001	.006	.05	.83	
Number of Caregivers	-.16	-.095	.033	8.29	.01	.40
Total Day Care Experience	.01	.003	.005	.26	.62	
Staff/Child Ratio	.32	1.566	.486	10.37	.005	
Staff/Child Ratio	.32	.585	.412	2.20	.17	.10
Education Level	.05	.005	.032	.02	.87	
Group Size	-.57	-.012	.004	8.7	.009	.34
Education Level	.05	-.012	.029	.2	.64	
Total Day Care Experience	.01	.003	.005	.20	.68	
Staff/Child Ratio	.32	.592	.427	1.9	.18	.10
Education Level	.05	.005	.033	.0	.88	
Total Day Care Experience	.01	-.001	.006	.0	.89	

Table A.6 (cont'd)

ADULT ACTIVITY
(N=21)

	r	B	SE_B	F	P_F	R^2
Group Size	.43	.007	.003	4.20	.05	.18
Number of Caregivers	-.16	- .017	.023	.52	.48	.03
Staff/Child Ratio	-.47	- .722	.308	5.48	.03	.22
Group Size	.43	.007	.003	4.16	.06	.19
Level of Education	-.01	.010	.026	.15	.70	
Group Size	.43	- .021	.024	.76	.39	.05
Day Care Experience	-.12	.004	.005	.51	.48	
Number of Caregivers	-.16	- .033	.028	1.34	.26	.28
Staff/Child Ratio	-.47	1.067	.426	6.25	.02	
Number of Caregivers	-.16	.028	.030	.92	.35	.29
Day Care Experience	.12	.003	.005	.38	.55	
Staff/Child Ratio	-.47	-1.046	.435	5.77	.03	
Staff/Child Ratio	-.47	- .536	.345	2.41	.14	.28
Group Size	.43	.004	.004	1.34	.26	
Staff/Child Ratio	-.47	- .756	.313	5.84	.03	.25
Day Care Experience	.12	.004	.004	.75	.40	
Group Size	.43	- .807	.003	3.72	.07	.19
Education Level	-.01	.010	.026	.15	.71	
Day Care Experience	.12	.001	.005	.06	.81	
Staff/Child Ratio	-.47	- .585	.360	2.64	.12	.29
Group Size	.43	.004	.004	.92	.35	
Day Care Experience	.10	.003	.005	.38	.55	

Table A.6 (cont'd)

CENTER ACTIVITY
(N=21)

	r	B	SE_B	F	P_F	R^2
Number of Caregivers	.40	.091	.048	3.63	.07	.16
Staff/Child Ratio	.36	1.225	.732	2.80	.11	.13
Level of Education	-.22					
Day Care Experience	.29	.014	.011	1.69	.21	.08
Number of Caregivers	.40	.088	.048	3.37	.08	.20
Level of Education	-.22	- .051	.056	.85	.37	
Number of Caregivers	.40	.077	.050	2.38	.14	.23
Day Care Experience		.010	.011	.78	.39	
Level of Education	-.22	- .050	.056	.87	.36	
Level of Education	-.22					
Day Care Experience	.29					
Staff/Child Ratio	.36	1.121	.731	2.35	.14	.19
Day Care Experience	.29	.012	.010	1.32	.27	
Group Size	-.09	- .007	.008	.71	.41	.16
Education Level	-.22	- .066	.060	1.23	.28	
Day Care Experience	.29	.015	.011	1.39	.26	
Staff/Child Ratio	.36	1.151	.730	2.49	.13	.24
Education Level	-.22	- .059	.056	1.11	.31	
Day Care Experience	.29	.012	.010	1.23	.28	

Table A.6 (cont'd)

OBSERVES
(N=21)

	r	B	SE_B	F	P_F	R^2
Group Size	.66	.026	.007	14.88	.001	.44
Staff/Child Ratio	-.40	-1.481	.787	3.54	.07	.16
Education Level	-.16	-.046	.065	.50	.51	.03
Group Size	.66	.026	.007	14.40	.002	.44
Total Day Care Experience	.03	-.004	.010	.18	.68	
Number of Caregivers	.02	.143	.067	4.57	.04	.33
Staff/Child Ratio	-.40	-2.984	1.007	8.77	.008	
Staff/Child Ratio	-.40	-1.463	.799	3.36	.08	.18
Education Level	.16	-.042	.061	.48	.50	
Staff/Child Ratio	-.40	-1.495	.826	3.28	.08	.18
Education Level	-.16	-.042	.063	.44	.52	
Total Day Care Experience	.03	-.004	.012	.10	.75	

Table A.6 (cont'd)

MANAGEMENT BEHAVIOR
(N=21)

	r	B	SE_B	F	P_F	R^2
Group Size	-.48	-.008	.003	5.62	.03	.23
Group Size	-.48	-.007	.003	5.07	.03	.38
Total Day Care Experience	-.50	-.010	.005	4.38	.05	
Group Size	-.48	-.010	.003	9.30	.007	.49
Total Day Care Experience	-.50	-.009	.005	3.80	.06	
Staff/Child Ratio	-.12	-.624	.322	3.75	.07	
Total Day Care Experience	-.50	-.012	.005	6.41	.02	.25
Number of Caregivers	-.48	-.041	.021	3.87	.06	.38
Total Day Care Experience	-.50	.010	.004	4.52	.04	
Group Size	-.48	-.007	.003	5.14	.03	.42
Total Day Care Experience	-.50	-.010	.004	5.88	.02	
Number of Caregivers	-.48	-.039	.020	3.93	.06	.48
Total Day Care Experience	-.50	-.009	.004	4.95	.04	
Education Level	.35	.040	.022	3.29	.08	
Number of Caregivers	-.48	-.041	.021	3.87	.06	.38
Total Day Care Experience	-.50	.010	.004	4.52	.04	
Staff/Child Ratio	-.12	.564	.32	3.07	.09	.51
Total Day Care Experience	-.50	-.009	.004	4.46	.05	
Group Size	-.48	-.010	.003	8.69	.009	
Group Size	-.48	-.006	.003	3.84	.06	.48
Education Level	.35	.032	.022	2.09	.16	
Total Day Care Experience	-.50	-.010	.004	6.23	.02	
Staff/Child Ratio	-.12	-.118	.317	.14	.71	.37
Education Level	.35	.042	.024	3.09	.09	
Total Day Care Experience	-.50	-.011	.005	6.16	.02	
Education Level	.35	.042	.024	3.20	.09	.36
Total Day Care Experience	-.50	-.011	.004	6.82	.02	

Table A.6 (cont'd)

SOCIAL INTERACTION
(N=21)

	r	B	SE_B	F	P_F	R^2
Group Size	-.76	-.021	.004	25.55	.000	.57
Staff/Child Ratio	.44	1.173	.552	4.51	.04	.19
Education Level	.30	.062	.045	1.87	.18	.09
Group Size	-.76	-.020	.004	22.12	.000	.59
Education Level	.30	.028	.032	.79	.61	
Staff/Child Ratio	.44	2.889	.557	26.00	.000	.60
Number of Caregivers	-.15	-.158	.037	18.41	.001	
Staff/Child Ratio	.44	2.753	.544	25.59	.002	.64
Number of Caregivers	-.15	-.152	.036	17.70	.001	
Education Level	.30	.432	.030	2.7	.17	
Staff/Child Ratio	.44	1.15	.54	4.54	.04	.27
Education Level	.30	-.059	.041	2.01	.17	
Group Size	-.76	-.020	.004	20.16	.001	.59
Education Level	.30	.028	.033	.75	.60	
Total Day Care Experience	-.16	-.002	.006	.10	.75	
Staff/Child Ratio	.44	1.217	.542	5.04	.04	.32
Education Level	.30	.057	.041	1.90	.18	
Total Day Care Experience	-.16	-.008	.008	1.06	.32	
Education Level	.30	.061	.046	1.75	.20	.11
Total Day Care Experience	-.16	-.006	.009	.46	.51	

Table A.7

STEM-AND-LEAF DISPLAYS OF ADULT FOCUS VARIABLES BY LEVEL OF EDUCATION

INSTRUCTS

AAT		HS	
.00-.03	2 0	.00-.03	1 0 3 2 3 3
.03-.06	6 5 5 5 4	.03-.64	4 5 3 5 3 4 5
.06-.09	7 6 8 7 6 8 9	.06-.09	9 7 7
.09-.12	0 9 2	.09-.12	2 1
.12-.15	4 2 4	.12-.15	3 4 3
.15-.18	8	.15-.18	
.18-.21	2	.18-.21	
.21-.24	2	.21-.24	
.24-.27		.24-.27	
.27-.30		.27-.30	9

ADULT ACTIVITY

AAT		HS	
.00-.02	0 0 0 0 1 1 0 0 0 0 0 1	.00-.02	1 1 0 2 1 0 0 0
.02-.04	4	.02-.04	3 4 3
.04-.06	4 4	.04-.06	4
.06-.08	6 7	.06-.08	8 7 6 7 7 6
.08-.10	0 9 0	.08-.10	8 0
.10-.12		.10-.12	
.12-.14	3	.12-.14	3
.14-.16		.14-.16	
.16-.18		.16-.18	
.18-.20		.18-.20	
.20-.22	0	.20-.22	0

Table A.7 (cont'd)

CENTER ACTIVITY

AAT

.00-.05	
.05-.10	
.10-.15	4 4
.15-.20	9
.20-.25	0 0
.25-.30	7 8 8 8 8 9
.30-.35	0
.35-.40	8 6 8
.40-.45	2 0 1 0
.45-.50	6 6 7
.50-.55	
.55-.60	
.60-.65	

HS

.00-.05	4
.05-.10	
.10- .15	
.15-.20	8
.20-.25	1 2
.25-.30	9 8 6 7
.30-.35	2 4 5 1 1 1
.35-.40	7
.40-.45	5 1 3 0
.45-.50	8
.50-.55	4
.55-.60	
.60-.65	4

OBSERVES

AAT

.00-.L0	8
.10-.20	7 9 6 4 2 3 7
.20-.30	2 3 8 4 5 2 3 3
.30-.40	6 8 4 0
.40-.50	6 2
.50-.60	
.60-.70	
.70-.80	

HS

.00-.10	
.10-.20	3 7 9 9
.20-.30	4 5 0 3 3 6 4 2
.30-.40	8 0 1 3 6 1
.40-.50	2 2
.50-.60	2
.60-.70	
.70-.80	8

Table A.7 (cont'd)

MANAGEMENT BEHAVIOR

AAT		HS	
.00-.05		.00-.05	4
.05-.10	8	.05-.10	9 6 8 7
.10-.15	1 2 0 3 4 2 0 2	.10- .15	4 2 5 1 2 0 2 0 3
.15-.20	0 7 9 6 0 7	.15-.20	6 6
.20-.25	3 2 0 2 1 1	.20-.25	0 0 0 0 1 1
.25-.30		.25-.30	
.30-.35	0	.30-.35	

SOCIAL INTERACTION

AAT		HS	
.00-.15		.00-.05	2
.05-.10		.05-.10	9 9 7 9 0
.10-.15	4 4 3	.10-.15	0 1 3 0 2 3
.15-.20	9 0 5 6 0	.15-.20	9 0 0
.20-.25	2 2 3 1 3 3	.20-.25	1 3 3 2 3
.30-.35	6 8 7 7 5 8	.30-.34	8
.35-.40	1	.35-.40	
.40-.45	8	.40-.45	
.45-50		.45-.50	8

Table A.8

STEM-AND-LEAF DISPLAYS OF ADULT FOCUS VARIABLES BY JOB ROLE

INSTRUCTS

Teachers

.00-.03	1 0 3 3
.03-.06	5 5 4 4 5 3 4 5
.06-.09	7 8 8 9 8 7
.09-.12	1 1 0 9 2 2 1
.12-.15	2 4 2 4 4
.15-.18	
.18-.21	
.21-.24	
.24-.27	
.27-.30	

Aides

.00-.03	2 1 0 3 2
.03-.06	6 5 5 4 6
.06-.09	7 6 7 6 9 7 8
.09-.12	
.12-.15	3 4
.15-.18	8
.18-.21	
.21-.24	2
.24-.27	
.27-.30	9

ADULT ACTIVITY

Teachers

.00-.02	0 0 1 1 0 1 0 0 0 0 1
.02-.14	4 3
.04-.06	4 4 6 4
.06-.08	8 7 6 6 7
.08-.10	0 9 8 8 0 9 3
.10-.12	
.12-.14	3 3
.14-.16	
.16-.18	
.18-.20	
.20-.22	

Aides

.00-.02	1 2 1 0 0 0 0 0 0 1 0 0
.02-.04	4 4 3
.04-.06	
.06-.08	7 7 7 6
.08-.10	0
.10-.12	
.12-.14	
.14-.16	
.16-.18	
.18-.20	
.20-.22	0 0

Table A.8 (cont'd)

CENTER ACTIVITY

Teachers

.00-.05	
.05-.10	
.10-.15	4
.15-.20	
.20-.25	1 2
.25-.30	7 8 9 6
.30-.35	1 1 1 0
.35-.40	8 7 6 6 8
.40-.45	1 0 1 3 0 4 0
.45-.50	7 6 7
.50-.55	4 3 4
.55-.60	
.60-.65	4

Aides

.00-.05	4
.05-.10	
.10-.15	4
.15-.20	8 9
.20-.25	0 0
.25-.30	7 8 8 9 8 8 6
.30-.35	2 4 5
.35-.40	
.40-.45	2 0 5
.45-.50	6 8
.50-.55	
.55-.60	
.50-.65	4

OBSERVES

Teachers

.00-.10	6 8
.10-.20	0 2 6 1 2 9 6 4 2 3 7 9 9
.20-.30	6 2 3 3 3 6 4
.30-.40	7 0 1 3 6 1
.40-.50	2 2
.50-.60	
.60-.70	
.70-.80	

Aides

.00-.10	
.10-.20	7 3 7
.20-.30	2 3 8 4 5 2 4 5 0 2
.30-.40	6 8 4 8 0
.40-.50	6 2
.50-.60	2
.60-.70	
.70-.80	8

Table A.8 (cont'd)

MANAGEMENT BEHAVIOR

Teachers		Aides	
.00-.05		.00-.05	4
.05-.10	5 8 9 5	.05-.10	6 8 7
.10-.15	1 4 1 2 2 5 1 2 1 1 2 1 1	.10-.15	3 2 4 2 0 0 2 3
.15-.20	0 7 6 5 9	.15-.20	6 0 7 6
.20-.25	3 0 4 1 2 0 0 2	.20-.25	0 1 1 1 0
.25-.30		.25-.30	
.30-.35		.30-.35	0

SOCIAL INTERACTION

Teachers		Aides	
.00-.05	1	.00-.05	2
.05-.10	9 9	.15-.10	7 9 0
.10-.15	2 2 4 1 0 3	.10-.15	3 4 0 2 3
.15-.20	0 8 9 0 5 9 0	.15-.20	6 0 0
.20-.25	2 1 2 2 1 3 3	.20-.25	3 3 1 3 2 3
.25-.30	8 6 8 7 7 8	.25-.30	8 5
.30-.35	3 1	.30-.35	
.35-.40		.35-.40	8
.40-.45		.40-.45	
.45-.50		.45-.50	8

Table A.9

SCHEMATIC PLOTS OF ADULT FOCUS VARIABLES

INSTRUCTS

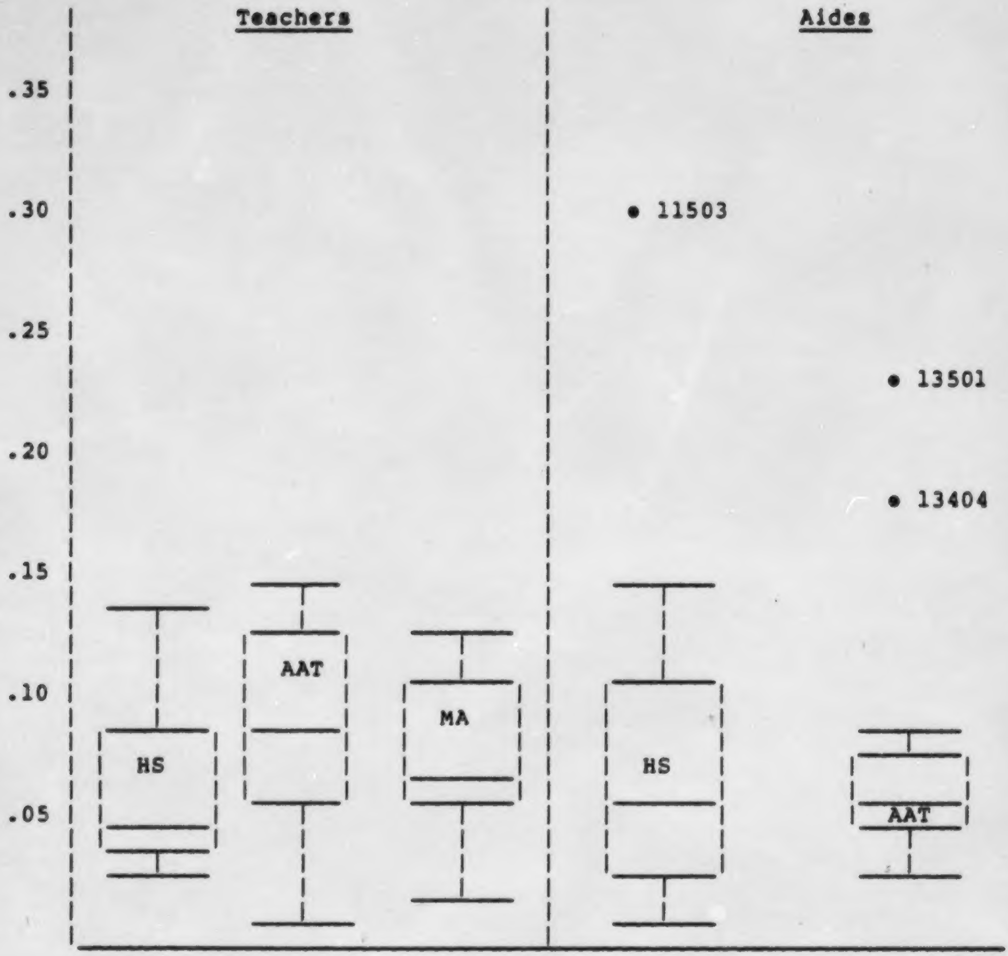


Table A.9 (cont'd)

ADULT ACTIVITY

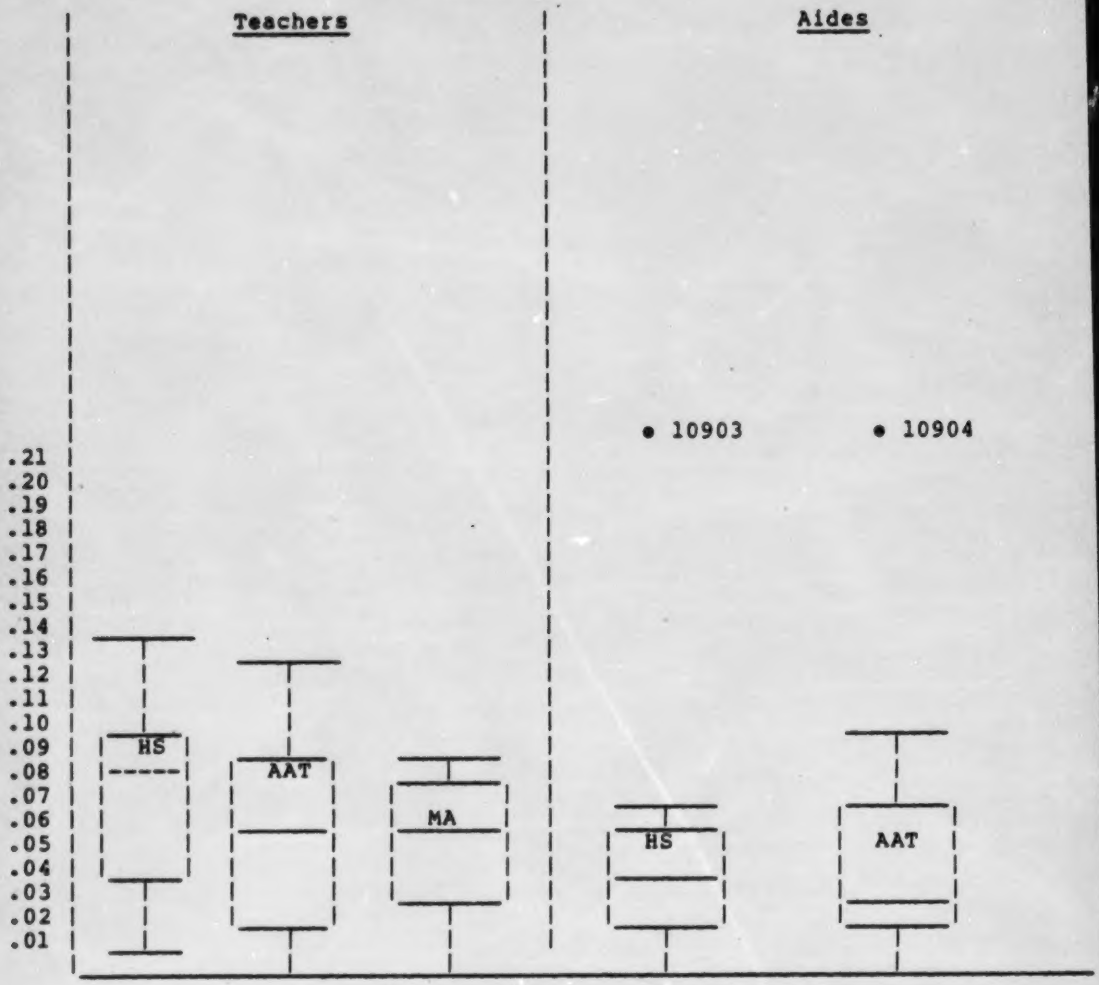
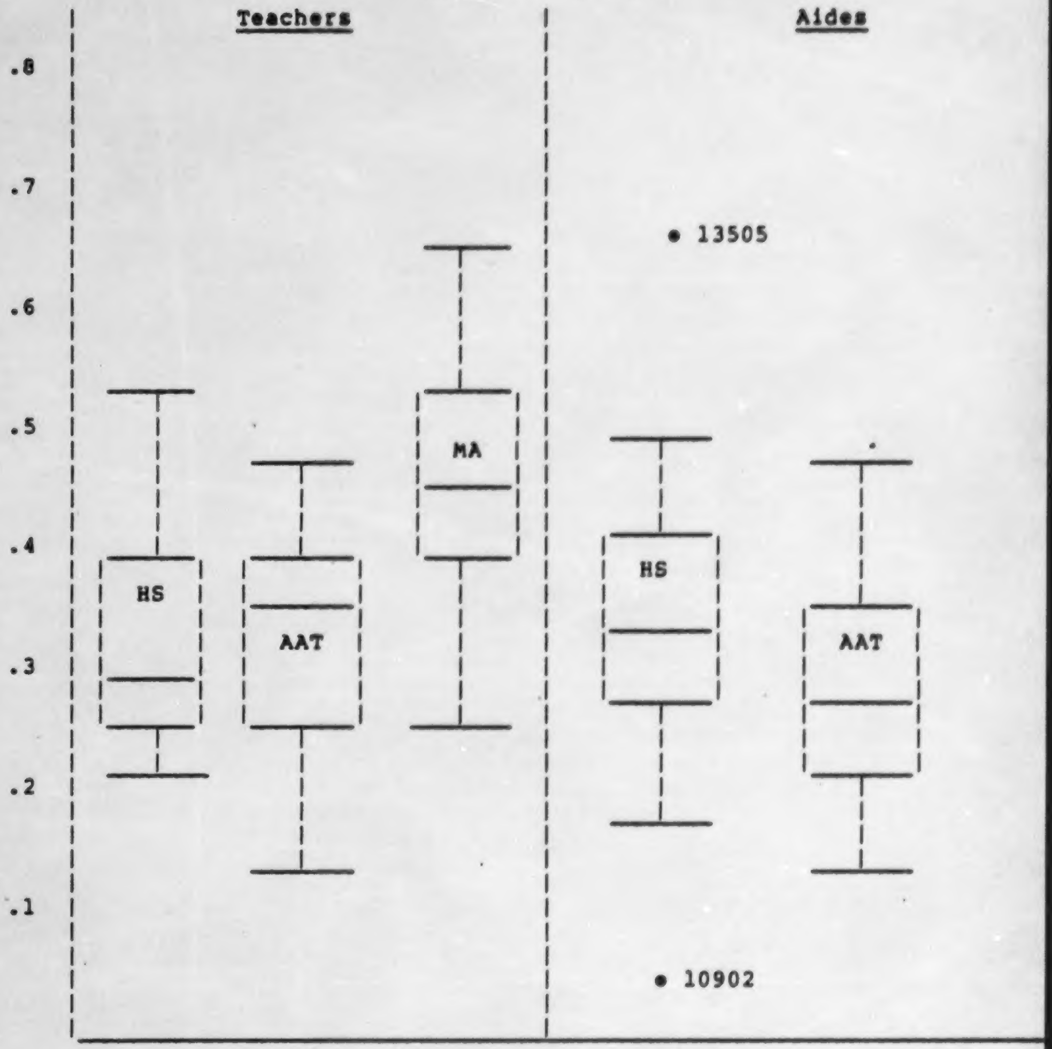


Table A.9 (cont'd)

CENTER ACTIVITY



531

Table A.9 (cont'd)

OBSERVES

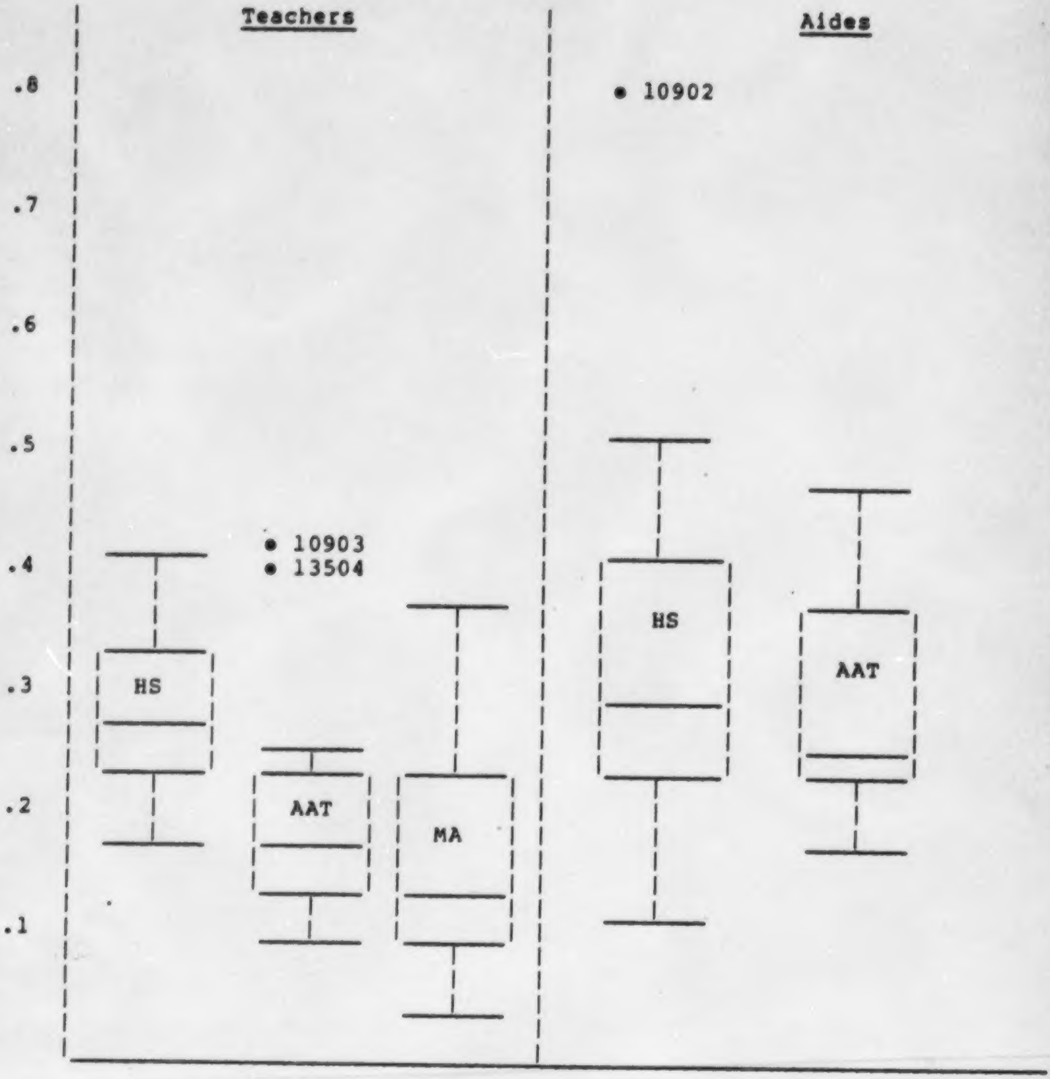


Table A.9 (cont'd)

MANAGEMENT BEHAVIOR

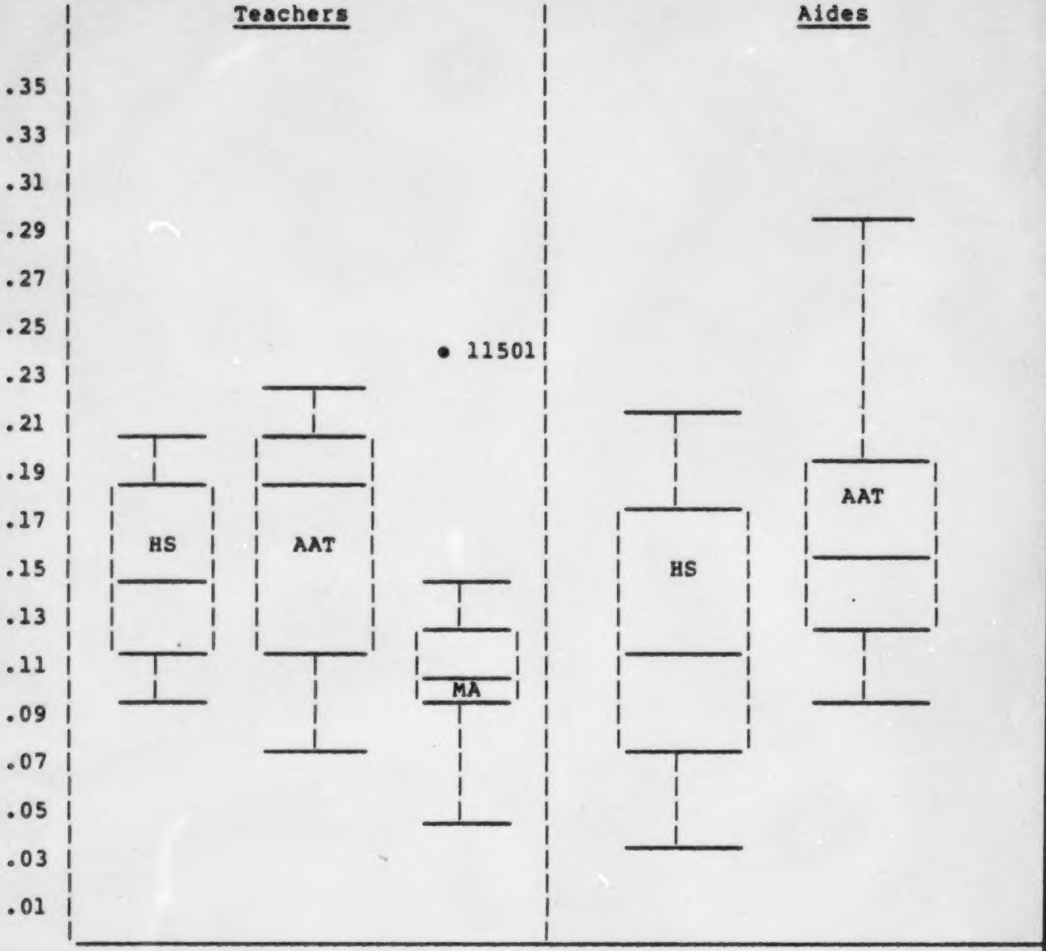


Table A.9 (cont'd)

SOCIAL INTERACTION

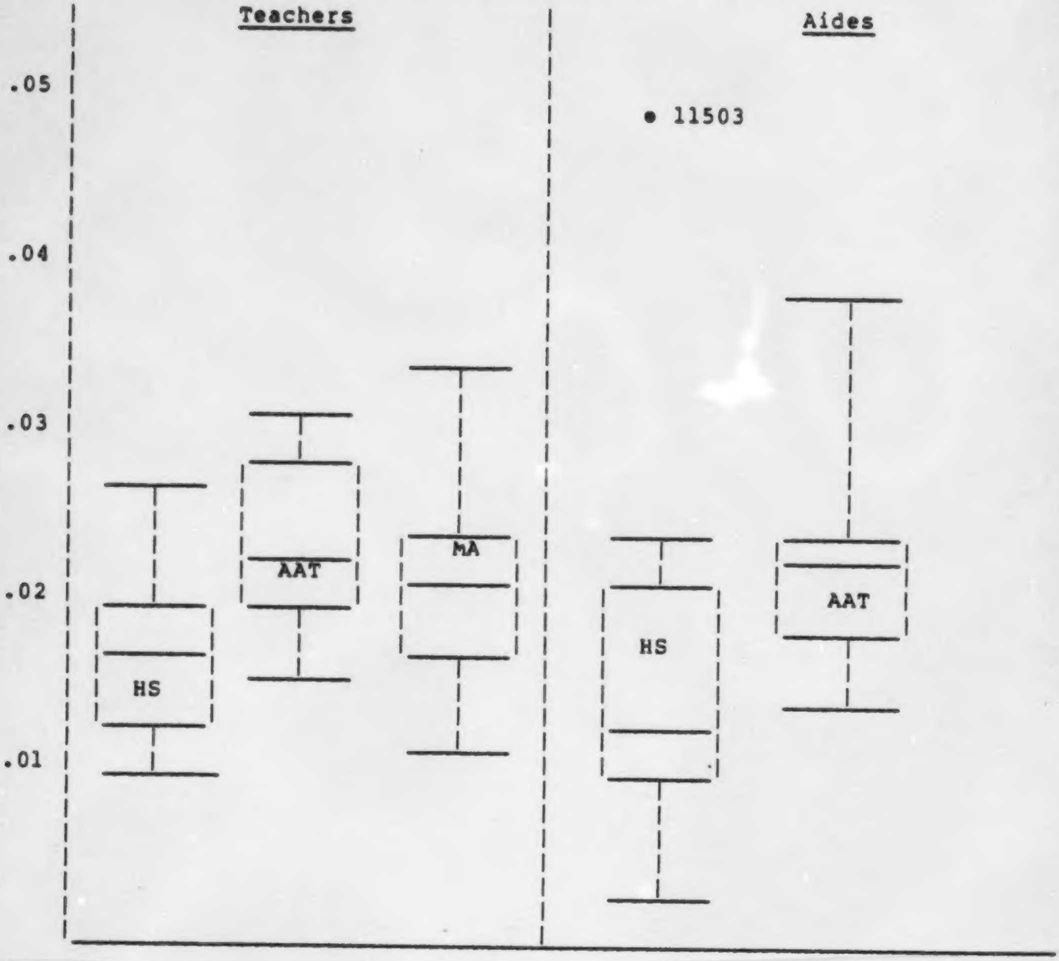


Table A.10

DIFFERENCES IN CAREGIVER BEHAVIORS BY LEVEL OF EDUCATION AND JOB ROLE

	MA's (n=8)		AAT's (n=11)		HS's (n=11)			
	Mean	S.D.	Mean	S.D.	Mean	S.D.	F	Sig.
Teachers								
Social Interaction	.21	.07	.23	.06	.17	.07	2.26	.12
Management Behavior	.12	.06	.17	.05	.15	.04	2.31	.12
Instructs	.08	.04	.09	.04	.06	.04	1.09	.35
Adult Activity	.04	.04	.04	.05	.06	.04	.39	.68
Center Activity	.46	.12	.35	.10	.34	.10	3.21	.06
Observes	.17	.10	.20	.11	.28	.07	3.17	.06
Aides			(n=11)		(n=11)			
Social Interaction	--	--	.22	.07	.16	.12	1.97	.16
Management Behavior	--	--	.17	.06	.13	.06	2.87	.07
Instructs	--	--	.08	.06	.08	.08	.01	.99
Adult Activity	--	--	.04	.06	.05	.06	.05	.95
Center Activity	--	--	.28	.10	.33	.16	.65	.53
Observes	--	--	.29	.09	.34	.19	.69	.51

Table A.11

PEARSON CORRELATIONS FOR ADULT FOCUS VARIABLES BY LEVEL OF EDUCATION AND JOB ROLES

	Instructs	Adult Activity	Center Activity	Observes	Management Behavior	Social Interaction
Group Size						
Teachers (n=30)			- .31	.40*		- .22
MA (n=8)		.69*				
AAT (n=11)	- .43	- .60		.54*		- .67*
HS (n=11)	.40		- .44		.35	
Aides (n=22)	- .54**	.44*		.67**	- .48*	- .75**
AAT (n=11)		.59*				- .36
HS (n=11)	- .82**	.40		.82**	- .58*	- .84**
Number of Caregivers						
Teachers (n=30)		- .36*				- .30*
MA (n=8)	- .59			.38		- .66*
AAT (n=11)	- .42	- .59*		.36		- .71**
HS (n=11)	.38					
Aides (n=22)			.38*		- .48**	
AAT (n=11)						
HS (n=11)	- .40		.57*		- .65*	
Staff/Child Ratio						
Teachers (n=30)		- .32*	.38*			
MA (n=8)	- .57	- .55				- .48
AAT (n=11)						
HS (n=11)			.59*	- .50		
Aides (n=22)	.31	- .48*	.37*	- .41*		.44*
AAT (n=11)	.43	- .42	.65*		- .48	
HS (n=11)	.79*	- .51		- .74*	.66*	.88**
Education Level						
Teachers (n=30)	.20		.40*	- .44**	- .19	.26
Aides (n=22)					.35*	.30
Total Day Care Experience						
Teachers (n=30)						
MA (n=8)	26					
AAT (n=11)			.37			
HS (n=11)			.31		- .42	
Aides (n=22)			.29		- .50**	
AAT (n=10)	.48				- .41	.31
HS (n=11)	- .34	.46	.48		- .64**	- .45
Previous Day Care Experience						
Teachers (n=29)	.19					.32*
MA (n=8)				- .53	.42	
AAT (n=11)			.48	- .55*		
HS (n=10)						.39
Aides (n=21)						
AAT (n=10)			- .39			
HS (n=11)		.57*	.49	- .42		
Center Experience						
Teachers (n=29)	.19					
MA (n=8)						
AAT (n=11)					- .35	
HS (n=10)						
Aides (n=21)					- .45*	- .27
AAT (n=10)					- .38	
HS (n=11)	- .34				- .67**	- .44

* p<.05
** p<.01

Correlations are reported for p<.15 only.

Table A.12

POLICY VARIABLE/CHILD FOCUS REGRESSION RESULTS
(N=30)

Dependent Variable	Independent Variable	r	B	SE_B	F	P_F	R^2
Monitors Environment	Group Size	.47	.004	.001	7.90	.01	.22
	Previous Day Care Experience	-.35	-.018	.009	3.83	.06	.12
	Previous Day Care Experience	-.35	-.024	.009	7.11	.01	.24
	Staff/Child Ratio	-.21	-.525	.251	4.37	.05	
Wanders	Previous Day Care Experience	-.52	-.018	.005	10.32	.01	.27
Gives Opinions	Group Size	-.38	-.002	.001	4.85	.04	.15
	Center Experience	-.37	-.006	.003	4.46	.04	.14
Gives Orders	Center Experience	.34	.001	.001	3.55	.07	.11
Receives Orders	Center Experience	-.38	-.004	.002	4.68	.04	.14
Receives Info, Help	Group Size	-.72	-.004	.001	29.89	.01	.52
	Number of Caregivers	-.49	-.020	.007	9.06	.01	.24
	Number of Caregivers	-.49	-.023	.007	13.68	.01	.34
	Staff/Child Ratio	.08	.318	.160	3.93	.06	
Receives General	Center Experience	-.46	-.005	.002	7.34	.01	.21
Part. In Open Activity	Center Experience	.55	.023	.007	12.40	.01	.31
	Center Experience	.55	.019	.0067	8.39	.01	.39
	Group Size	.45	.0034	.0017	3.79	.06	
Part. In Structured Activity	Group Size	-.38	-.006	.003	4.72	.04	.14

Table A.13

POLICY VARIABLES/CHILD FOCUS REGRESSION RESULTS
TEACHER-DIRECTED ACTIVITIES
(N=30)

Dependent Variable	Independent Variable	r	B	SE_B	F	P_F	R^2
Monitors Environment	Group Size	.44	.004	.001	6.80	.01	.20
	Number of Caregivers	.40	-.024	.011	5.18	.03	.16
Wanders	Group Size	.46	-.003	.001	7.33	.01	.21
	Number of Caregivers	.40	.020	.009	5.19	.03	.16
	Group Size	.46	.004	.001	10.01	.01	.30
	Staff/Child Ratio	.19	.356	.193	3.40	.08	
Gives Opinions	Group Size	-.32	-.002	.001	3.12	.09	.10
Gives Orders	Group Size	.32	.002	.001	3.20	.08	.10
Receives Orders	Previous Day Care Experience	.33	.006	.003	3.49	.07	.11
	Previous Day Care Experience	.33	.006	.003	4.41	.05	.21
	Level of Education	-.29	-.006	.003	3.45	.07	
Receives Info, Help	Previous Day Care Experience	.53	.035	.010	11.15	.01	.28
	Previous Day Care Experience	.53	-.027	.010	7.06	.01	.42
	Group Size	.52	-.004	.002	6.36	.02	
Receives General	Center Experience	-.31	-.003	.002	2.96	.10	.10
Part. In Open Activity	Previous Day Care Experience	-.35	-.014	.007	3.89	.06	.12
Part. In Structured Activity	Group Size	-.41	-.009	.004	5.70	.02	.17

Table A.14

POLICY VARIABLES/CHILD FOCUS REGRESSION RESULTS
FREE-PLAY SITUATIONS
(N=30)

Dependent Variable	Independent Variable	r	B	SE_B	F	P_F	R^2
Monitors Environment	Group Size	.31	.003	.002	2.94	.10	.10
	Previous Day Care Experience	-.30	-.016	.010	2.75	.11	.09
Wanders	Level of Education	-.35	-.015	.007	4.02	.05	.13
Gives Opinions	Number of Caregivers	-.40	-.019	.008	5.34	.03	.16
	Number of Caregivers	-.40	-.016	.008	3.94	.06	.25
	Center Experience	-.38	.008	.004	3.34	.08	
Gives Orders	Center Experience	.29	.002	.001	2.56	.12	.08
Receives Orders	Group Size	-.47	.002	.001	8.24	.01	.23
	Center Experience	-.35	.005	.002	3.90	.06	.12
Receives Info, Help	Group Size	-.32	-.001	.001	3.12	.09	.10
Receives General	Center Experience	-.41	-.006	.002	5.67	.02	.17
Part. In Open Activity	Center Experience	.55	.043	.012	12.44	.01	.31
Part. In Structured Activity	Center Experience	-.38	-.028	.013	4.85	.04	.15

Table A.15

POLICY VARIABLE/CDA REGRESSION RESULTS

(N=29)

Dependent Variable	Policy Variable	r	b	SE_B	F	P_F	R^2
Maintains Safe Classroom	Staff/Child Ratio	.30	.849	.512	2.75	.11	.09
Maintains Sanitary Classroom	Staff/Child Ratio	.39	.685	.308	4.95	.03	.15
	Staff/Child Ratio	.39	.706	.286	6.10	.02	.30
	Previous Day Care Experience	.37	.024	.010	5.41	.03	
Provides Gross Motor Toys	Previous Day Care Experience	.50	.066	.022	9.17	.005	.25
Provides Creative Play Mat.	Staff/Child Ratio	.49	2.482	.847	8.59	.007	.24
Arranges Classroom Well	Previous Day Care Experience	.43	.099	.040	6.03	.02	.18
	Previous Day Care Experience	.43	.101	.037	7.33	.01	.32
	Staff/Child Ratio	.49	2.332	1.027	5.16	.03	
Encourages Safety	Level of Education	.32	.104	.059	3.09	.09	.10
Manages Class Activities Well	Previous Day Care Experience	.58	.077	.021	13.56	.001	.33
	Previous Day Care Experience	.58	.085	.019	20.31	.00	.49
	Level of Education	.33	.090	.032	7.89	.009	
Encourages Active Play	Previous Day Care Experience	.62	.104	.025	16.86	.001	.38
Encourages Cog/Lang. Develop.	Group Size	-.56	-.030	.008	12.30	.002	.31
Encourages Good Self-Concept	Previous Day Care Experience	.46	.059	.022	7.27	.01	.21
	Previous Day Care Experience	.46	.060	.020	8.63	.007	.33
	Staff/Child Ratio	.33	1.209	.560	4.66	.04	
Encourages Self-Help	Center Experience	.31	.027	.016	2.96	.09	.10
	Center Experience	.31	.029	.015	3.51	.07	.18
	Staff/Child Ratio	.26	1.026	.656	2.45	.13	
	Center Experience	.31	.032	.015	4.72	.047	.29
	Staff/Child Ratio	.26	1.432	.660	4.75	.04	
	Level of Education	-.20	-.079	.041	3.81	.06	
Encourages Social Behavior	Previous Day Care Experience	.54	.111	.033	11.41	.002	.30
Constructs							
Physical Environment	Staff/Child Ratio	.45	.748	.287	6.80	.01	.20
Resources	Previous Day Care Experience	.41	.062	.027	5.36	.03	.17
	Previous Day Care Experience	.41	.064	.024	6.84	.01	.33
	Staff/Child Ratio	.40	1.701	.669	6.46	.02	
Classroom Management	Previous Day Care Experience	.51	.068	.022	9.37	.05	.26
	Previous Day Care Experience	.51	.076	.020	14.66	.001	.43
	Level of Education	.36	.096	.034	8.03	.009	
Child Orientation	Previous Day Care Experience	.48	.064	.023	8.10	.008	.23
	Previous Day Care Experience	.48	.066	.020	10.52	.003	.40
	Staff/Child Ratio	.39	1.512	.559	7.32	.01	
CDA Rating	Previous Day Care Experience	.52	.056	.018	9.87	.004	.27
	Previous Day Care Experience	.52	.057	.015	13.77	.001	.47
	Staff/Child Ratio	.43	1.320	.422	9.77	.004	

Table A.16

PEARSON CORRELATIONS FOR ADULT FOCUS AND CHILD FOCUS VARIABLES TEACHERS--TEACHER-DIRECTED AND FREE-PLAY COMBINED

	Instructs	Adult Activity	Center Activity	Observes	Management Behavior	Social Interaction
Monitors Env. (n=30)		.24		.20		-.27
MA (n=8)		.90**			-.62*	
AAT (n=11)			-.38			
Non-AAT (n=11)				.43	.35	-.36
Part. In Group-Open						
MA		.49	.53			-.53
AAT		-.60*		.37		
Non-AAT						
Part. In Group-Closed		-.21				.20
MA		-.70*	-.68*		.65*	.45
AAT						
Non-AAT	.44					.40
Does Nothing; Wanders				.24		-.38*
MA	-.51	.70*	.68*		-.59	-.73*
AAT						
Non-AAT						-.47
Gives Opinions	.27	.42**	-.21			.31*
MA	.65*	-.44	-.62*		.64*	.55
AAT		.69**		-.41		.52*
Non-AAT		.40				
Gives Orders			.29	-.19	-.28	
MA		.65*	.42		-.64*	
AAT						
Non-AAT			.46		-.52*	-.39
Receives Orders			-.25			
MA		-.85**	-.49		.78**	.72*
AAT						
Non-AAT	.44	.41			-.34	
Receives Info, Help	.24	-.24			.26	.37*
MA		-.47				
AAT					.57*	.50
Non-AAT						.34
Receives General		.25	-.22			
MA						
AAT		.65**			-.47	
Non-AAT	-.48		-.35		.44	

* p<.05
** p<.01
Correlations are reported for p<.15 only.

Table A.17

PEARSON CORRELATIONS FOR ADULT FOCUS AND CHILD FOCUS VARIABLES AIDES--TEACHER-DIRECTED AND FREE-PLAY COMBINED

		Instructs	Adult Activity	Center Activity	Observers	Management Behavior	Social Interaction
Monitors Env.	(n=22)	-.27		.26	.25	-.23	-.57
MA	(n=11)						-.61
Non-AAT	(n=11)	-.49			.38		-.53*
Part. In Group-Open		-.29	.30		.37*	-.66**	-.53**
AAT			.36			-.65	
Non-AAT		-.49			.48	-.55*	-.57*
Part. In Group-Closed		.43*		-.44*		.45*	.70**
AAT				-.71		.41	.68
Non-AAT		.71**			-.43	.36	.81**
Does Nothing; Wanders			.72**			-.25	-.24
AAT			.79				-.43
Non-AAT			.67**				
Gives Opinions					-.26	.53**	
AAT		-.42		-.47	.34	.46	
Non-AAT					-.38	.51*	
Gives Orders							
AAT		.54					
Non-AAT		-.39					
Receives Orders			-.42*			.40*	.25
AAT			-.70**				.39
Non-AAT						.53*	
Receives Info, Help		.40*	-.69**	.32	-.39*		.41*
AAT			-.78				.41
Non-AAT		.46	-.64*	.44	-.62*	.38	.48
Receives General			.36	-.24			
AAT		-.59					
Non-AAT			.50	-.49		.42	

* p<.05
** p<.01
Correlations are reported for p<.15 only.

Table A.18

PEARSON CORRELATIONS FOR ADULT FOCUS AND CHILD FOCUS VARIABLES TEACHERS--TEACHER-DIRECTED ACTIVITIES

		Instructs	Adult Activity	Center Activity	Observers	Management Behavior	Social Interaction
Monitors Env.	(n=30)		.31*			-.19	-.37*
MA	(n=8)		.81**			-.65*	
AAT	(n=11)				.40		-.39
Non-AAT	(n=11)						-.44
Part. In Group-Open			-.31*		-.39*	.41**	.28
MA					-.45	.76**	.61
AAT				.38		.51*	-.38
Non-AAT			-.44	.34	-.50	.38	
Part. In Group-Closed			-.37*				
MA			-.60				
AAT			-.38			.42	
Non-AAT		.40					.35
Does Nothing; Wanders					.19		-.31*
MA			.81**	.44		-.61	-.69*
AAT						-.36	
Non-AAT							-.39
Gives Opinions			.37*				
MA						-.44	
AAT			.52*	-.54*			.52*
Non-AAT			.45			-.59*	-.52*
Gives Orders				.23		-.27	
MA			.67*				
AAT			-.39			.40	
Non-AAT				.36		-.78**	-.46
Receives Orders							
MA		.42	-.46	-.67*		.67*	.60
AAT				.62*		-.44	
Non-AAT		.39					
Receives Info, Help			-.37*		-.21	.26	.31*
MA			-.49				
AAT							
Non-AAT			-.44				.42
Receives General							
MA							
AAT			.60*				.35
Non-AAT		-.67**	.35				-.48

* p<.05
** p<.01
Correlations are reported for p<.15 only.

Table A.19

PEARSON CORRELATIONS FOR ADULT FOCUS AND CHILD FOCUS VARIABLES AIDES--TEACHER-DIRECTED ACTIVITIES

	Instructs	Adult Activity	Center Activity	Observers	Management Behavior	Social Interaction
Monitors Env. (n=22)	-.35			.29		-.53**
AAT (n=11)						-.64*
Non-AAT (n=11)	-.65*			.57*		-.61*
Part. In Group-Open		-.30		-.23	.54**	.33
AAT					.77**	
Non-AAT	.72**	-.47		-.48		.75**
Part. In Group-Closed	.59**	-.34		-.33	.35	.72**
AAT	.63*		-.52*			.83**
Non-AAT	.63*	-.36		-.61*	.48	.72**
Does Nothing; Wanders	-.35	.78**		.28	-.35	-.36*
AAT	-.37	.82**				-.50
Non-AAT	-.35	.78**		.47	-.51*	-.39
Gives Opinions					.32	
AAT					-.45	-.46
Non-AAT	.35				.64*	
Gives Orders						-.25
AAT			.39			-.45
Non-AAT	-.36					
Receives Orders						
AAT	-.37		-.44	.41		
Non-AAT						
Receives Info, Help		-.58**	.59**	-.32		
AAT		-.71**				.41
Non-AAT		-.58*	.71**	-.52*		
Receives General		.34				
AAT	-.56*		.42			-.41
Non-AAT		.68**	-.49		.38	

* p<.05
** p<.01
Correlations are reported for p<.15 only.

Table A.20

PEARSON CORRELATIONS FOR ADULT FOCUS AND CHILD FOCUS VARIABLES TEACHERS--FREE PLAY

		Instructs	Adult Activity	Center Activity	Observers	Management Behavior	Social Interaction
Monitors Env.	(n=30)		.24	-.23	.21		
MA	(n=8)		.83**			-.56	
AAT	(n=11)			-.68**			
Non-AAT	(n=11)						
Part. In Group-Open			-.21				-.19
MA		-.46		.65*			-.60
AAT			-.68**		.42		-.38
Non-AAT						-.47	
Part. In Group-Closed							
MA			-.62*	-.81**		.76**	.54
AAT			.46	.43	-.43		
Non-AAT		.34		-.41		.60*	
Does Nothing; Wanders							-.33*
MA		-.80**		.56			-.44
AAT		.40	.59*				
Non-AAT							
Gives Opinions				-.28		.22	.33*
MA		.42		-.54		.72*	.50
AAT			.73**				.58*
Non-AAT				-.35		.43	
Gives Orders				.27		-.23	
MA			.55	.43		-.72*	
AAT							
Non-AAT				.42		-.38	
Receives Orders		.22	.25	-.21			
MA			-.59				.50
AAT		.47	.38	-.43		.44	.50
Non-AAT			.36				
Receives Info, Help		.35*					.32*
MA		.47					.48
AAT			.40	-.37			.45
Non-AAT		.54*					
Receives General			.27	-.36*	.21		
MA				-.80**	.85**		
AAT			.65**				
Non-AAT		-.35		-.36		.60*	

* p<.05
** p<.01
Correlations are reported for p<.15 only.

Table A.21

PEARSON CORRELATIONS FOR ADULT FOCUS AND CHILD FOCUS VARIABLES AIDES--FREE PLAY

	Instructs	Adult Activity	Center Activity	Observers	Management Behavior	Social Interaction
Monitors Envr. (n=22)			.24			-.37*
AAT (n=11)		.35	.42			-.43
Non-AAT (n=11)						
Part. In Group-Open	-.28	.24	.25	.29	-.69**	-.45*
AAT		.40			-.73**	
Non-AAT	-.48				-.58*	-.51
Part. In Group-Closed	.23		-.47**		.41*	.52**
AAT			-.64*		.62*	.52*
Non-AAT	.43					.52*
Does Nothing; Wanders		.30				
AAT						
Non-AAT						
Gives Opinions					.39*	
AAT			-.38		.56*	
Non-AAT						
Gives Orders						
AAT	.37					
Non-AAT				.36		
Receives Orders	.42*	-.43*		-.34	.59**	.36*
AAT		-.81**			.43	
Non-AAT	.50			-.39	.73**	.38
Receives Info, Help	.38*	-.38*		-.43*	.74**	.33
AAT				-.42	.85**	
Non-AAT	.62*	-.60*		-.54*	.61*	.57*
Receives General		.44*	-.27			
AAT	-.38	.56*			-.56	
Non-AAT		.40			.39	

* p<.05
** p<.01
Correlations are reported for p<.15 only.

Table A.22

ADULT FOCUS/PSI GAIN SCORE REGRESSION RESULTS

	Teachers (N=29)					
	r	B	SE_B	F	P_F	R^2
Constructs						
Observes	-.32	-7.624	4.372	3.04	.09	.10
Observes	-.32	-9.451	4.637	4.15	.05	.14
Social Interaction	-.08	-8.242	7.263	1.29	.27	
Pol. Vars. & Constructs						
Group Size	-.37	-.315	.071	19.69	.000	.60
Staff/Child Ratio	.56	18.721	10.539	3.16	.08	
Social Interaction	-.08	-14.129	5.361	6.94	.01	
Observes	-.32	-7.173	3.339	4.61	.04	

	Aides (N=22)					
	r	B	SE_B	F	P_F	R^2
Constructs						
Social Interaction	.49	13.109	5.242	6.25	.02	.24
Social Interaction	.49	10.181	5.165	3.90	.06	.36
Adult Activity	.48	-14.663	7.650	3.67	.07	
Pol. Vars. & Constructs						
Group Size	-.66	-.557	.198	7.95	.01	.59
Staff/Child Ratio	.59	20.723	10.465	3.92	.06	
Social Interaction	.49	-3.397	6.291	.29	.60	
Observes		5.622	3.473	2.62	.12	

Table A.23

ADULT FOCUS/PSI GAIN SCORE REGRESSION RESULTS

Teachers (N=29)

	r	B	SE_B	F	P_F	R^2
Constructs						
Management Behavior	-.29	-18.430	11.765	2.45	.13	.08
Management Behavior	-.29	-25.19	11.71	4.63	.04	.20
Adult Activity	-.24	-30.31	15.47	3.84	.06	
Negative	-.36	-81.18	44.05	3.40	.07	.21
Positive	.32	59.45	36.24	2.69	.11	
Pol. Vars. & Constructs						
Previous Day Care Experience	.50	1.847	.698	6.99	.01	.38
Management Behavior	-.29	-22.417	10.611	4.46	.04	
Adult Activity	-.24	21.049	14.379	2.14	.15	
Previous Day Care Experience	.50	1.444	.802	3.24	.08	.41
Group Size	-.23	-.108	.135	.64	.56	
Staff/Child Ratio	.07	-18.837	20.497	.84	.63	
Management Behavior	-.29	-23.554	10.841	4.72	.04	
Adult Activity	-.24	-20.391	15.167	1.81	.19	

Aides (N=21)

	r	B	SE_B	F	P_F	R^2
Constructs						
Social Interaction	.30	11.938	8.596	1.93	.17	.09
Pol. Vars. & Constructs						
Previous Day Care Experience	.36	1.752	.984	3.17	.09	.23
Social Interaction	.30	12.195	8.145	2.24	.15	
Previous Day Care Experience	.36	2.212	1.052	4.42	.05	.31
Group Size	-.19	.203	.325	.31	.54	
Staff/Child Ratio	.01	-22.429	20.977	1.14	.30	
Social Interaction	.30	22.899	12.818	3.19	.09	

Table A.24

CHILD FOCUS/COGNITIVE GAIN SCORE REGRESSION RESULTS
COMBINATION SITUATIONS
(N=30)

Dependent Variable	Independent Variable	r	B	SE_B	F	P_F	R^2
PSI Gain Score	Receives Information	.46	33.890	15.020	5.09	.03	.15
	Receives Information	.46	31.370	14.872	4.45	.04	.21
	Gives Opinions	.19	29.096	20.685	1.98	.17	
	Group Size	-.38	- .314	.076	17.08	.00056	.48
	Staff/Child Ratio	.57	22.211	11.437	3.77	.060	
	Receives Orders	.10	- 30.078	30.742	1.30	.26	
PPVT Gain Score	Wanders	-.49	- 60.421	19.925	9.20	.0053	.25
	Wanders	-.49	- 59.135	18.754	9.94	.0042	.36
	Gives Orders	.28	253.100	117.517	4.64	.038	
	Wanders	-.49	- 53.124	16.740	10.07	.0041	.51
	Gives Orders	.28	334.098	107.772	9.61	.0048	
	Structured Activity	.50	23.411	8.071	8.41	.0074	
	Previous Day Care Experience	.51	1.434	.678	4.47	.0042	.59
	Gives Orders	.28	324.477	101.334	10.25	.0039	
	Structured Activity	.50	22.933	7.584	9.14	.0058	
	Wanders	-.49	- 33.754	18.200	3.44	.072	

Table A.25

CHILD FOCUS/COGNITIVE GAIN SCORE REGRESSION RESULTS
TEACHER-DIRECTED ACTIVITIES
(N=30)

Dependent Variable	Independent Variable	r	B	SE_B	F	P_F	R^2
PSI Gain Score	Gives Opinions	.57	43.515	15.018	8.40	.0072	.23
	Gives Opinions	.57	46.325	14.210	10.63	.003	.34
	Part. in Structured Activity	.25	8.545	4.008	4.55	.04	
	Gives Opinions	.57	50.620	14.280	12.57	.002	.39
	Part. in Structured Activity	.25	7.223	4.044	3.19	.08	
	Receives Information	.26	9.320	6.578	2.01	.17	
	Gives Opinions	.57	49.756	13.582	13.42	.001	.47
	Part. in Structured Activity	.25	6.372	3.870	2.71	.11	
	Receives Information	.26	13.403	16.599	4.12	.05	
	Receives Orders	-.22	-49.572	25.555	3.76	.06	
	Group Size	-.30	- .231	.071	10.71	.003	.58
	Staff/Child Ratio	.067	21.932	10.491	4.37	.04	
	Gives Opinions	.57	24.900	12.666	3.86	.06	
	Receives Orders	-.22	-36.712	21.710	2.86	.10	
PPVT Gain Score	Part. in Open Activity	-.34	-50.661	21.639	5.48	.02	.16
	Part. in Open Activity	-.34	-43.965	21.404	4.22	.05	.24
	Wanders	-.36	-28.396	17.294	2.70	.11	
	Previous Day Care Experience	.55	1.973	.745	7.03	.01	.33
	Structured Activity	.44	10.616	6.381	2.77	.10	
	Previous Day Care Experience	.51	1.993	.812	6.02	.02	.28
	Wanders	-.36	-16.983	18.094	0.88	.64	

Table A.26

CHILD FOCUS/COGNITIVE GAIN SCORE REGRESSION RESULTS
FREE-PLAY SITUATIONS
(N=30)

Dependent Variable	Independent Variable	r	B	SE_B	F	P_F	R^2
PSI Gain Score	Receives Orders	.33	52.708	27.221	3.75	.06	.12
	Receives Orders	.33	54.729	26.636	4.22	.05	.19
	Monitors Environment	-.21	- 14.633	9.615	2.32	.14	
	Group Size		- .302	.073	16.99	.001	.48
	Staff/Child Ratio		22.874	11.491	3.96	.05	
	Gives Orders	.17	49.998	43.106	1.35	.25	
PPVT Gain Scores	Gives Orders	.40	204.645	75.750	7.30	.01	.21
	Gives Orders	.40	194.982	70.356	7.68	.01	.34
	Monitors Environment	-.42	- 30.878	13.085	5.57	.024	
	Gives Orders	.40	175.000	64.909	7.27	.012	.47
	Monitors Environment	-.42	- 31.352	11.981	6.85	.014	
	Wanders	-.33	- 37.326	14.974	6.21	.018	
	Gives Orders	.40	167.870	61.892	7.36	.011	.54
	Monitors Environment	-.42	- 36.436	11.706	9.69	.005	
	Wanders	-.33	- 30.836	14.647	4.43	.043	
	Receives Information	.32	47.830	24.868	3.70	.063	
	Previous Day Care Experience	.51	2.726	.611	19.93	.003	.57
	Gives Orders	.40	233.987	59.301	15.57	.008	
	Structured Activity	.33	15.203	6.248	5.92	.02	
	Previous Day Care Experience	.51	1.798	.609	8.72	.007	.60
	Gives Orders	.40	187.937	57.531	10.67	.003	
	Monitors Environment	-.42	- 26.933	11.483	5.50	.026	
	Receives Information	.32	49.009	22.900	4.60	.040	

Table A.27

CDA/PSI GAIN SCORE REGRESSION RESULTS
(N=30)

Independent Variable	r	B	SE_B	F	P_F	R^2
CDA Rating	.31	4.736	2.713	3.05	.09	.10
Child Orientation	.19	2.303	2.199	1.10	.30	.04
Encourages Active Play	.17	1.897	2.058	.85	.63	.03
Encourages Cog/Lang. Dev.	.26	2.277	1.574	2.09	.16	.07
Encourages Good Self-Concept	.11	1.379	2.303	.36	.56	.01
Encourages Self Help	.23	2.968	2.398	1.53	.22	.05
Encourages Social Behavior	.00	(Not computable)				
Classroom Management	.28	4.012	2.567	2.44	.13	.08
Encourages Safety	.28	2.537	1.653	2.36	.13	.08
Manages Class Activities Well	.24	3.509	2.638	1.77	.19	.06
Resources	.38	4.298	1.953	4.84	.03	.15
Provides Gross Motor Materials	.30	4.128	2.476	2.78	.10	.09
Provides Creative Play Materials	.30	4.510	1.574	8.20	.008	.23
Arranges Classroom Well	.22	.154	1.30	1.40	.25	.05
Physical Environment	.03	.744	5.262	.02	.88	.00
Maintains Safe Classroom	-.01	(Not computable)				
Maintains Sanitary Classroom	.05	1.149	4.754	.06	.80	.00

Table A.28

CDA/PPVT REGRESSION RESULTS
(N=30)

Independent Variable	r	B	SE_B	F	P_F	R^2
CDA Rating	.52	12.130	3.743	10.50	.003	.27
Child Orientation	.48	8.571	2.995	8.19	.008	.23
Encourages Active Play	.38	5.420	2.969	3.33	.07	.11
Encourages Cog/Lang. Dev.	.42	5.470	2.246	5.93	.02	1.7
Encourages Good Self-Concept	.49	9.045	3.039	8.86	.006	.24
Encourages Self Help	.053	1.061	3.795	.08	.76	.003
Encourages Social Behavior	.51	6.178	1.963	9.90	.004	.26
Classroom Management	.42	9.026	3.711	5.92	.02	.17
Encourages Safety	.12	1.706	2.576	.44	.52	.01
Manages Class Activities Well	.49	10.786	3.622	8.87	.006	.24
Resources	.52	8.967	2.798	10.28	.004	.27
Provides Gross Motor Materials	.48	10.067	3.470	8.42	.007	.23
Provides Creative Play Materials	.39	5.702	2.553	4.99	.03	.15
Arranges Classroom Well	.50	5.360	1.774	9.12	.005	.25
Physical Environment	.14	6.125	8.014	.58	.54	.02
Maintains Safe Classroom	-.11	-2.95	4.98	.35	.57	.012
Maintains Sanitary Classroom	.30	11.544	6.822	2.86	.10	.09